AD_____
(Leave blank)

Award Number: **W81XWH-10-1-0556**

TITLE: MicroRNA gene regulatory networks in peripheral nerve sheath tumors

PRINCIPAL INVESTIGATOR:

Subbaya Subramanian, PhD

CONTRACTING ORGANIZATION:

University of Minnesota
Minneapolis-55455

REPORT DATE: September 2013

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
September 2013	Final	15 August 2010 - 14 August 2013

4. TITLE AND SUBTITLE

MicroRNA gene regulatory networks in peripheral nerve sheath tumors

5a. CONTRACT NUMBER
W81XWH-10-1-0556

5b. GRANT NUMBER
W81XWH-10-1-0556

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Subbaya Subramanian subree@umn.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Minnesota
Office of the Sponsored projects
200 Oak street SE, Minneapolis-55455

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT: One of the main hurdles to develop effective therapy for malignant peripheral nerve sheath tumors (MPNST) is the lack of understanding of molecular mechanisms regulating cancer genes and metastasis. Oncogenic epithelial-to-mesenchymal transition (EMT) is a critical step for metastasis and is closely related with transcriptional changes of many key genes involved in cell polarity, cell-cell adhesion, and cell migration. Multiple abnormal signaling in this pathways and processes can initiate, promote, and maintain EMT process. To understand the molecular networks that regulates EMT pathway and promote metastasis in MPNST, we systematically analyzed the microRNA (miRNA)–mediated gene regulatory networks using miRNA and MRNA expression profiles generated from normal Schwan cells, MPNSTs tumor tissues and cell lines. Both negative- and positive- correlation networks based on gene and miRNA expression data were generated. These correlations were further mined by referring to various relational database (protein-protein interactions, canonical pathways, transcription factor-to-target prediction) and genomics data (copy-number alteration, differential DNA methylation). We identified six candidate network modules, which potentially control the preferential activation of TGF-beta/SMAD signaling to TGF-beta/non-SMAD signaling and the induction of cancer cell stemness. By applying different levels of data integration and exploration, we could identify several units which take part in EMT of MPNST. Reconstructed networks suggests that miRNAs actively participate in transcription control of cancer genes and cause aberrant modification of core pathways responsible for transformation and metastasis. MPNST development.

15. SUBJECT TERMS
MPNST, microRNAs, miR-29 knockout

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	55	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

Introduction……. 3

Body…… 3

Key Research Accomplishments……..… 9

Reportable Outcomes…… 9

Conclusion…… 10

Final report summary……...10

References…… 11

Appendices…… 11

Introduction

MPNST are aggressive sarcomas often associated with NF1 gene inactivating mutations. MPNSTs develop due to malignant transformation of pre-existing benign dermal or plexiform neurofibromas, with the latter having more chances of malignant transformation. Development of MPNSTs from neurofibromas is a complex process. Several studies have found differential expression of genes between benign and malignant tumors suggesting the role of several genes and pathways in transformation. Such comparative studies alone could not attribute a 'cause or effect' relationship of these genes and pathways. So, detailed mechanisms of malignant transformation still remain to be understood.

Based on comparative gene/miRNAs expression studies using tumors of various stages and cell lines of neurofibromas MPNSTs and Schwann cells we have determined various genes and miRNAs that could be involved in malignant transformation. Since miRNAs can regulate multiple genes simultaneously, deregulation of miRNAs leads to wide spread miss-regulation of multiple genes as in the case of key genes operating high in the order of cellular homeostasis pathways. We have observed miR29 has been significantly downregulated in MPNSTs compared to neurofibromas. Several studies have implicated the role of miR29 in activating p53 and DNMTs, which are major players in tumorigenesis, progression and transformation. We have also noticed the deregulation of miRNAS such as miR34a and miR214, whose role in MPNSTs has yet to be fully established. Hence, we hypothesize that deregulation of 'miRNAs gene regulatory networks' play a significant role in malignant transformation of MPNSTs.

Body

Our objectives in this study is to determine the gene/miRNAs signatures of MPNST susceptibility which could serve as predictive biomarkers and functionally characterize their role in malignant transformation of neurofibromas into MPNSTs using both in *vitro* and *in vivo* mouse models which might lead to miRNAs based therapeutics. We proposed to carry out this work with specific objectives listed below.

Specific Aims:

Aim # 1: Define profiles of MPNST susceptibility based on microRNA and gene expression signatures in human and mouse tumors. To accomplish this, we will:

1.1. Generate miRNA profiles for benign and malignant PNSTs from human and mouse tumors and for human PNST cell lines.

1.2. Analyze gene expression profiles (mRNA) for human and mouse PNSTs and human PNST cell lines.

1.3. Validate the expression profiles of candidate miRNAs and mRNAs by quantitative PCR.

Aim # 2: Decipher the regulatory pathways mediated by microRNAs that are etiologically significant for malignant transformation to MPNST. To achieve this, we will:

2.1. Identify putative miRNA regulatory networks by in *silico* analysis of miRNA and mRNA expression data. We will analyze regulatory interactions involving miRNAs that are differentially expressed in PNSTs as observed in our preliminary studies (for example, miR-29, miR-214 and miR-34).

2.2. Perform functional validation of miRNA- mRNA association *in vitro* using human PNST cell lines and understand the biological consequences of miRNA modulation.

2.3. Develop novel miRNA based diagnostic markers of MPNSTs using tissue microarrays.

Aim # 3: Engineer miR-29a/b dysregulation *in vivo* in the setting of benign neurofibroma. To accomplish this, we will:

3.1. Develop a mouse model that shows 'accelerated malignant transformation' phenotype with Schwann cell/Schwann cell precursor specific (Dhh-Cre) disruption of miR-29a/b.

3.2. Characterize the phenotypic and genotypic features of these engineered mice.

3.3. Identify predictive biomarkers of malignant transformation using blood plasma obtained from engineered mice that have potential to show accelerated malignant transformation phenotype.

In the past years we have completed the following tasks

Completed tasks 2011

1. All the necessary, Institutional review board, Institutional Biosafety board and Animal use approval were obtained to conduct the proposed experiments.

2. We have collected the benign and malignant peripheral nerve sheath tumor tissues from mice and have analyzed the microRNA expression profiles.

3. MicroRNA and gene expression profiles have been completed form the human normal and tumor tissues. We have used normal Schwann cells for normalization of the data. The miRNA expression data is published and publically available in www.oncomiR.umn.edu

4. miRNA- gene network analysis is currently being carried out and we have identified potential candidate miRNAs and genes for further functional analysis.

5. Experiments are currently being standardized for miRNA transfection and functional studies.

6. We have also started the construction of miR-29 knockout vectors.

Completed tasks 2012

1. Completed generation of knockout constructs for both miR-29a/b1 and miR-29 b2/c

2. DNA electroporation and screening of positive clones.

3. Identified SUZ12P as a potential regulator of microRNAs and driver of malignant transformation through integrated genomic approaches

This annual report covers the work carried out from Aug 15th 2012 to Aug 15th 2013. The following tasks were approved to complete during this period.

Completed tasks 2013:

1. Analysis of microRNA gene regulatory networks in MPNSTs.
2. Develop miR-29a-b knockout constructs.
3. Development of miR-29 knockout mice, task is on going due to difficulties in obtaining the knocked out ES clones.
4. Analysis of circulating miRNA markers in MPNSTs

1. Generating knock-out mice of miR29 family:

miR29 family miRNA are distributed in 2 clusters in mice. miR29a and miR29b1 are on 6th chromosome and miR29b2/miR29c are on 1st chromosome. We don't have conclusive evidences to confirm whether these miRNAs are expressed in clusters or regulated individually. In our experiments we have analyzed expression levels of all 4 miRNAs of miR29 family and found that their levels are different. However it could be due to differential rate of degradation of these miRNAs. Therefore we don't have conclusive evidences to confirm co-expression or co-regulation of these miRNAs. Although these miRNAs share significant sequence similarity they have the potential to target different target genes. However this does not rule out the possibility of redundancy in their functions. So it is essential to knock out all four miRNA29 family miRNAs individually and also in combinations to study their role in tumor transformation / growth.

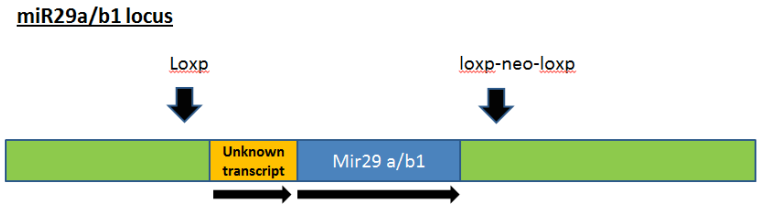


Figure 1a: miR29a/b1 locus with known transcripts indicated in black arrows. The sites for lox p insertions are also shown.

Designing Gene targeting vector for miR29 family.

Since we did not have any conclusive information regarding the transcription start site, promoters of miR29a and miR29b1, it is very challenging to decide the sequences to be knocked out to get rid of miR29s specifically without affecting neighboring transcripts. It is also difficult to ensure normal expression of these miRNAs after introducing the lox-p sites for conditional deletion of alleles. So, we analyzed the genomic locus around the miRNAs, all the available transcripts available and also RNA seq data of Argounaute1. Argonaute2 and argonaute3 clip seq data (ref; starbase). We found several uncharectorised transcripts and noncoding RNAs in this locus. After aligning all the sequences we narrowed down to regions which are free of any transcripts and/or safe to insert the lox-p sites.

Knock-out construct of miR29ab1:

miR29a/miR29b1 are located on 6th chromosome. We dont have any conclusive evidence to confirm the coexpression and/or co-regulation of these two miRNAs. Levels of expression of these two miRNAs are usually not similar although they are located on the same chromosome very close to each other. We have observed several transcripts reported from this locus other than miRNAs and their putative precursors. There was one long transcript (#1 in figure 1A) which includes miR29a/b1 miRNAs, but there is a second transcript (#2 in figure1A) upstream to the former and ends almost at the beginning of the same transcript. Hence it was difficult to predict the transcription start site of these miRNAs. Further these multiple transcripts restrict the choice of sites for lox-p insertions. We considered all such factors in designing out gene targeting vectors.

In order to construct a gene targeting vector for miR29a/b1 we designed a new vector with puc19 backbone with a specially designed multiple cloning site consisting of several restrictions sites. miR29a/b1 locus (of approximately 13kb) was amplified in 3 fragments. On either side of fragment 2 which consists of miR29a/b1 we inserted lox-p and loxp-neo-loxp cassette (figure 1B). a schematic map of the gene targeting vector is in figure 1B. since the promoter sites/regulatory sequences of miRNAs was not known we inserted loxp-neo-loxp sites downstream to fragment 2.

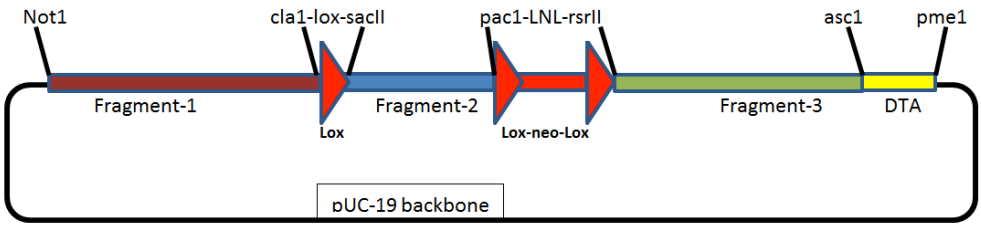


Figure 1b: Map of gene targeting vector designed for miR29a/b1 locus.

Gene targeting vector for miR29b2 and miR29c:

miR29b2 and miR29c are located on chromosome 1. Similar to miR29a and b1 we don't have any evidences to confirm whether they are expressed and/or coregulated. Analyzing the transcripts from this locus revealed several overlapping transcripts in both directions, making it very difficult to choose sites for lox-p sequence insertions

A schematic of the transcripts is shown in **figure2A.** miR29b2 and C seems to be a part of a very long non-coding RNA. There are several other smaller transcripts reported from this locus as well in the same orientation (fig 2A). They could be either degraded products of long noncoding RNAs or independent smaller transcripts. There are two other small transcripts in apposite orientation to miR29b2 and c. considering the possible transcriptions start sites of such transcripts and promoter sites we
could narrow down to 2 very small regions where we can insert loxp sites (to determine the locus for deletion) as indicated in figure 2A and B. loxp-neo-loxp cassette was inserted after the 2nd fragment to avoid possible leaky transcription/elongation of the unknown transcript and gene expressed in apposite orientation as indicated in the fig 2A. loxp site was inserted upstream of miRNAs, since precise locations of promoters are not known it was the safest approach.

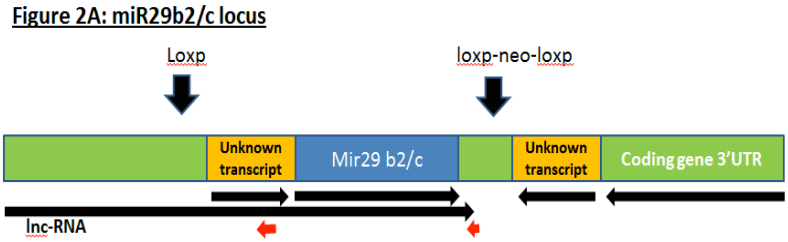


Figure 2a: miR29b2/c locus with known transcripts indicated in black arrows. The sites for lox p insertions are also shown.

A schematic map of the designed gene targeting vector is shown in figure 2B. Both the vectors designed for conditional deletion of included regions. Even the choices of loxp and loxp-neo-loxp cassette were chosen based on the direction of miRNA transcription and the neighboring genes and non-coding RNAs. In our view this is the best possible way to delete the miR29 family miRNAs without/least affecting other neighboring genes and transcripts.

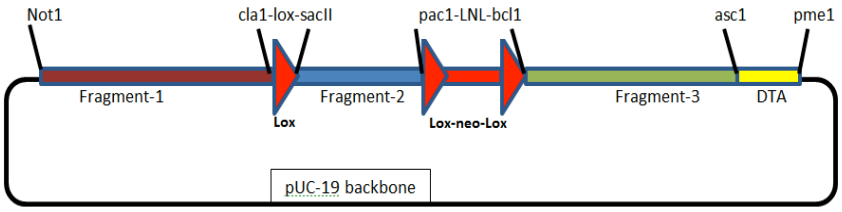


Figure 2b: Map of gene targeting vector designed for miR29b2/c locus.

Gene targeting in mouse ES cells:
In order to do gene targeting with both miR29a/b1 and miR29b2/c vectors, we linearized both gene targeting vectors and transfected them into C57BL/6J ES cells individually. Currently we are screening selected clones for site specific integration of gene targeting vectors.

We were disappointed that we failed to develop a miR-29 family knockout mice even after 3 consecutive attempts. We screened hundreds of colonies and none of the colonies seems to contain the correct targeting. We are currently redesigning the constructs to remove the DTA selection to maximize the colony numbers that may increase the chance of get a correct targeted clone.

II. Role of a pseudogene in transformation of MPNST
In order to investigate the mechanisms of transformation of tumors people have tried to explore the role of gene, miRNAs, CNVs etc. most of them are studied in human cell lines and animal models where it is possible to establish cause and effect relationships. However there could be many human specific elements in causing this transformation. here we have identified one such factor which might play a role in transformation of neurofibroma into MPNST.

Carefull analysis of nf1 locus which is often deleted/mutated in MPNSTs revealed that there are some hotspots between which microdeletions are usually observed. microdeletions are often seen in suz12 gene and a pseudogene of suz12 called suz12p (figure 3A). although suz12 and suz12p share more than 88% similarity in genomic regions suz12p transcripts are very small (~700nts) and share very little sequences (of ~500bases) with suz12 gene (fig3b). This facilitates intra-chromosomal recombination leading to micro-deletion of locus between them. The deleted locus contains many genes, miRNA and pseudogenes including a tumor suppressor *Nf1* [1] and many other genes, miRNA and pseudogenes. suz12 is a component of polycomb repressor complex (PRC2) which plays very critical role in cell differentiation and development.

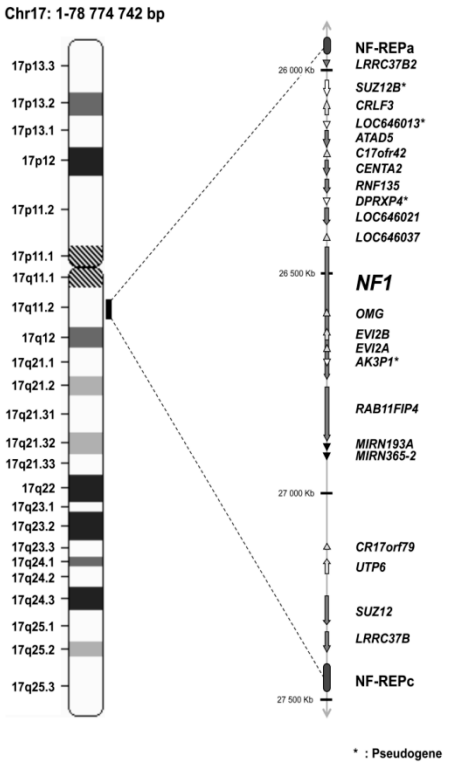


Figure 3: Schematic diagram showing NF1 locus; * indicates pseudogenes.

These polycomb genes are very critical in early developmental stages of an organism and highly conserved across species. They execute their function through *Polycomb Repressor Complex* (PRC1/2/3) which induce methylation of histones (markers of inactive chromatin) leading to suppression of genes. PRC are essential to maintain stemness of a cell and have to be suppressed to undergo differentiation during early development [2]. But in most of the cancers these genes are overexpressed implicating a de-differentiated state of tumor/transforming cells or at least tumor initiating cells [2, 3]. But their mechanisms need to be understood.
Pseudogenes are often found in most of the species and some are very specific to humans (ex:suz12p). Depending on the location of the pseudogene they might play a serious role in genomic recombination leading to chromosomal aberrations like deletions, recombination etc. many such chromosomal aberrations are associated with various cancers. Often pseudogenes are transcribed but the transcripts could be different from the homologous functional gene itself (at least partially). Their expression is spatially and temporally regulated. Therefore it is possible that they play significant role in regulating expression of various genes and noncoding RNAs and might be involved in other pathways including RNAi and Chromatin Modifications.

Our preliminary analysis of PAR-Clip data of Ago1 and Ago3 revealed *SUZ12P* transcripts are associated with RNAi machinery (see ref [4, 5]) implicating their role in RNAi mediated regulation. *Overall SUZ12P pseudogene and transcripts have the potential to play important role in gene/genome regulation.*
In our efforts to establish a correlation with cancer, we have observed that *SUZ12P* is overexpressed in MPNSTs compared to benign tumors. We observed a similar trend in Schwann cells and MPNST cell lines as well. **Figure 4**.

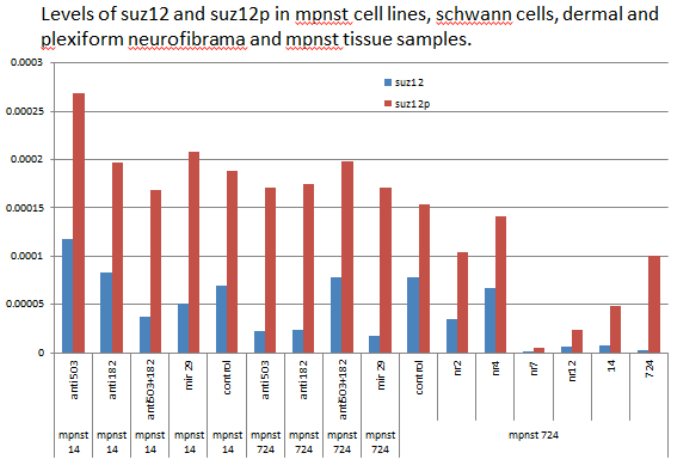
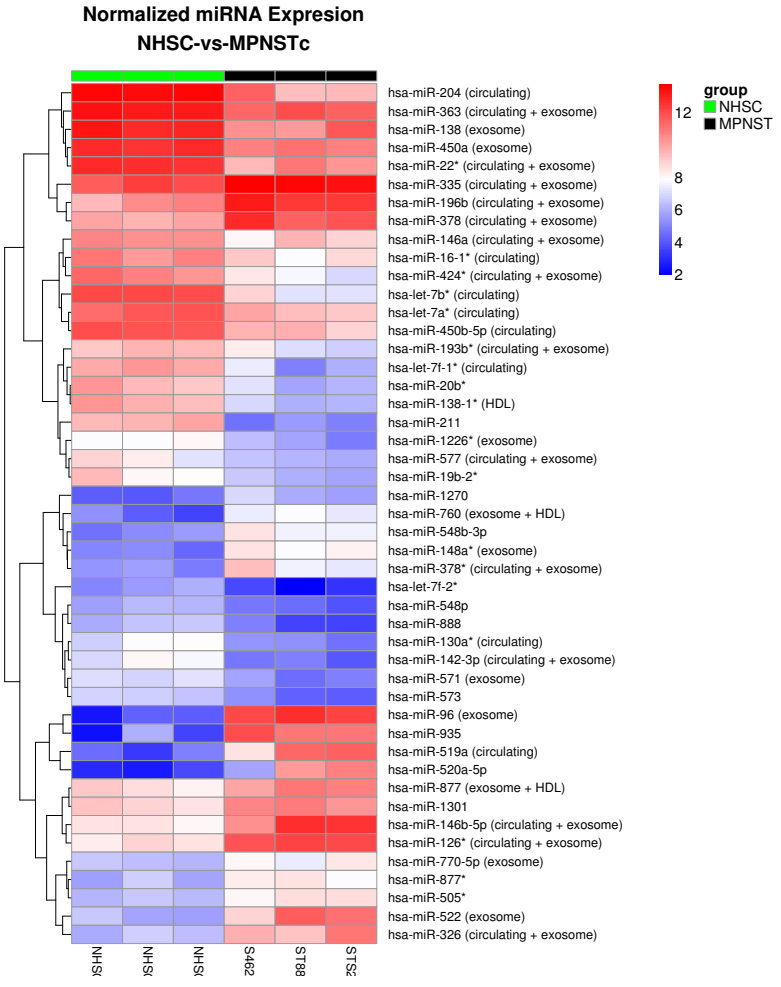


Fig4: figure shows the levels of suz12 and suz12p in schwann cells (nr2), dermal neurofibroma (nr4), plexiform neurofibroma (nr7) and MPNST (nr12) tissues and cells lines mpnst 14 (14), and mpnst 724 (724). Mpnst 14 and 724 cells were transfected with mimics of miR29, and antimirs of mir182 and miR503 individually and in

We have predicted and observed that several microRNAs can target SUZ12 &/or SUZ12P. Unlike commonly used predictions which consider only 3'UTRs of genes, we included coding regions, 5'UTRs and 3'UTRs avoiding any bias. Some of the predicted miRNAs include miR200c (involved in EMT, tumor transformation and metastasis) [6, 7, 8, 9], miR302, (inducing pluripotency) [10, 11] and many other important miRNAs like miR200b, 203, 214, 182, 21, 29, 520, 503 etc., An interesting outcome of our prediction is that miR-503, 182 and 520 may target 5'UTR, coding region and promoter region of SUZ12 respectively, and all of these can target SUZ12P transcripts. We have observed that overexpression of miR29 and miR503 leads to suppression of SUZ12P **(figure 4)**. Hence we think SUZ12P can regulate the levels of SUZ12 by titrating-out/ quenching of miRNAs that can suppress SUZ12 gene. This kind of buffering/titration is a novel mechanism of tuning the expression levels of genes by competing for the same regulatory RNAs. In a broader perspective, SUZ12P when overexpressed in cancer cells can thus release the suppression of many other genes inhibited by miR200b/200c/302/503/182/214 etc, which are known to be involved in EMT, pluripotency and other pathways driving tumor progression and transformation. If confirmed this mechanism would be a first of its kind where a pseudogene SUZ12P regulates polycomb mediated pathways, RNAi mediated pathways & gene/miRNA/lnc-RNAs expression contributing to tumor transformation.

In order to confirm the role of suz12 and suz12p in tumor transformation we would like to perform both gain-of-function and loss of function studies of both these gene/pseudogene. So we have constructed shRNA constructs to knock-down the expression of suz12/suz12p individually and both together and also full length transcripts to overexpress both genes. Our vectors have a luciferase reporter in an independent cassette to facilitate in-vivo tracking of cells in xenograft models. We would like to investigate the effect of loss and gain of suz12/suz12p on cell proliferation, viability, migration and invasion. We found Transient overexpression/ suppression of suz12/suz12p in cells showed only mild effects so currently we are selecting stably integrated cells of MPNST with all overexpression and shRNA constructs.

In order to confirm whether the role of suz12p is through RNA or some unknown translated product, we have designed suz12p constructs fused to T7 RNA polymerase promoter to produce RNA by in-vitro transcription. The in vitro transcribed RNA was transfected into mpnst cells and Cells were harvested to collect RNA and protein. Currently we are analyzing the expression levels of various genes/miRNAs and cellular parameters like viability, invasion, proliferation and migration.



III. MicroRNAs expressed in MPNST and relevance to circulating miRNAs in exosome.

We carried out analysis of miRNA expression profiles and compared to the circulating miRNAs. The heat map shows differentially expressed miRs in the controlling network (based on our array data with background correction + quantile normalization). This is an ongoing study, we have obtained potential candidate miRNAs which will be

tested in the serum samples obtained from the MPNST patients.

Fig 5: Heat map representation of circulating miRNAs relevant to MPNSTs.

IV. MicroRNA gene regulatory Network analysis.

One of the main hurdles to develop effective therapy for malignant peripheral nerve sheath tumors (MPNST) is the lack of understanding of molecular mechanisms regulating cancer genes and metastasis. Oncogenic epithelial-to-mesenchymal transition (EMT) is a critical step for metastasis and is closely related with transcriptional changes of many key genes involved in cell polarity, cell-cell adhesion, and cell migration. Multiple abnormal signaling in this pathways and processes can initiate, promote, and maintain EMT process. To understand the molecular networks that regulates EMT pathway and promote metastasis in MPNST, we systematically analyzed the microRNA (miRNA)–mediated gene regulatory networks using miRNA and MRNA expression profiles generated from normal Schwan cells, MPNSTs tumor tissues and cell lines.

Both negative- and positive- correlation networks based on gene and miRNA expression data were generated. These correlations were further mined by referring to various relational database (protein-protein interactions, canonical pathways, transcription factor-to-target prediction) and genomics data (copy-number alteration, differential DNA methylation). We identified six candidate network modules, which potentially control the preferential activation of TGF-beta/SMAD signaling to TGF-beta/non-SMAD signaling and the induction of cancer cell stemness. By applying different levels of data integration and exploration, we could identify several units which take part in EMT of MPNST. Reconstructed networks from this study suggests that miRNAs actively participate in transcription control of cancer genes and cause aberrant modification of core pathways responsible for transformation and metastasis in MPNST development. (See attached draft manuscript for details).

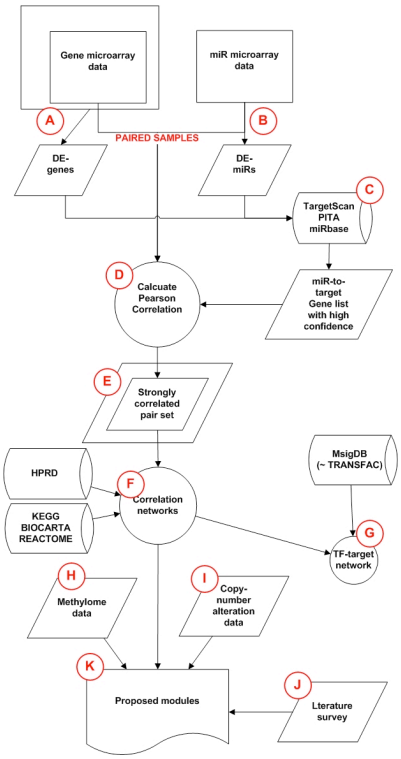


Fig 6: Overall analysis workflow.

V. Differential target selection of miRNAs.

MicroRNAs (miRNAs) are important players of post-transcriptional gene regulation. Individual miRNAs can target multiple mRNAs and a single mRNA can be targeted by many miRNAs. We hypothesized that miRNAs select and regulate their targets based on their own expression levels, those of their target mRNAs and triggered feedback loops. We studied the effects of varying concentrations of let-7a-7f and the miR-17-92 cluster plasmids on the reporter genes carrying either DICE R- or cMYC -3'UTR in Huh-7 cells. We showed that let-7 significantly downregulated expression of DICE R 3'UTR reporter at lower concentrations, but selectively downregulated expression of a cMYC 3'UTR reporter at higher dose. This miRNA dose-dependent target selection was also confirmed in other target genes, including CC ND1, CDKN1 and E2F1. After overexpressing let-7a-7f or the miR-17-92 clusters at wide-ranging doses, the target genes displayed a nonlinear correlation to the transfected miRNA. Further, by comparing the expression levels of let-7a and miR-17-5p, along with their selected target genes in 3 different cell lines, we found that the knockdown dose of each miRNA was directly related to their baseline expression level, that of the target gene and feedback loops. These findings were supported by gene modulation studies using endogenous levels of miR-29, -1 and -206 and a luciferase reporter system in multiple cell lines. Finally, we determined that the miR-17-92 cluster affected cell viability in a dose-dependent manner. In conclusion, we have shown that miRNAs potentially select their targets in a dose-dependent and nonlinear fashion that affects biological function; and this represents a novel mechanism by which miRNAs orchestrate the finely tuned balance of cell function. (See reference 12 and attached original article for details).

Key research accomplishments:

1. Constructed miRNA gene regulatory network in MPNST

2. Identified potential circulating miRNA relevant to MPNSTs.

3. Identified SUZ12p as a potential regulator of microRNAs and driver of malignant transformation through integrated genomic approaches.

4. Determined the biology of target gene selection in cancer cells.

Reportable outcomes 2013:
We have published a paper on genes that can function as competing endogenous RNA that can affect the microRNA networks. We have also developed and tested hypothesis that target gene selection is dependent on the dose of miRNA at any given time. A manuscript is under preparation that details various miRNA networks in MPNSTs.

Final report summary.
The collection of tumor tissue materials were carried out with appropriate Institutional review board and Institutional Biosafety board approvals. We also obtained the necessary animal use approval to conduct the DoD approved mice experiments. We completed miRNA expression profiles of over 25 MPNST human and mouse tissues. The miRNA expression data is published and publically available in www.oncomiR.umn.edu Subsequently, a comprehensive bioinformatics analysis was carried out to decipher the miRNA- gene networks in MPNSTs. Based on the analysis we have identified potential candidate miRNAs and genes for further functional analysis. Several experiments for miRNA functional analysis and gene networks were standardized. We completed construction of 6 knockout vectors for miR-29 family members that include miR-29a, 29b and 29c. These knockout vectors were used in the DNA electroporation and screening of positive ES clones. We had difficult in screening the clones as we obtained many false positives and changed the strategy for screening of true positives. We identified SUZ12P as a potential regulator of microRNAs and driver of malignant transformation through integrated genomic approaches. Functional analysis of Suz12p revealed that its expression levels could significantly affect it protein coding functional counter part SUZ12. We also carried out extensive analysis of circulating miRNA markers in MPNSTs. The manuscript is currently under preparation. We published the following 5 papers during the course of this project.

Publications
A research and review article was published in this project period. These articles are attached in appendices.

1. Sarver A, Subramanian S. Competitive endogenous RNA Database. *Bioinformation* 2012 8(15): 731-733.
2. Subramanian S and Kartha RV. MicroRNA mediated gene regulations in human sarcomas. *Cell Mol Life Sci 2012* 69: 3571-3585
3. Shu J, Xia Z, Li L, Liang ET, Slipek N, Shen D, Foo J, Subramanian S, Steer C. Dose-dependent differential mRNA target selection and regulation by let-7a-7f and miR-17-92 cluster microRNAs. *RNA Biol 2012* 9:1275-1287.
4. Choi K, Jegga AG, Hajeri P, Subramanian S, Ratner N MicroRNA mediated Gene Regulatory Networks in Malignant Peripheral Nerve Sheath Tumors (manuscript in preparation)
5. Subramanian S and Kartha RV. MicroRNA mediated gene regulations in human sarcomas. *Cell Mol Life Sci 2012* Nov;69(21):3571-85

Conclusions:
With this funding support we have generated extensive data in our understanding of miRNA mediated gene regulation in MPNSTs. We also generated a database of competing endogenous RNAs that will also be applicable to study other cancer types. We also created a network of miRNA gene interaction map and showed that the differential target gene section is cancer cells are based on the expression levels of miRNAs.

References:

1. Holzel, M.; Huang, S.; Koster, J.; Ora, I.; Lakeman, A.; Caron, H.; Nijkamp, W.; Xie, J.; Callens, T.; Asgharzadeh, S.; Seeger, R.C.; Messiaen, L.; Versteeg, R.; Bernards, R. NF1 is a tumor suppressor in neuroblastoma that determines retinoic acid response and disease outcome. *Cell* **2010**, *142*, 218-229.
2. Gieni, R.S.; Hendzel, M.J. Polycomb group protein gene silencing, non-coding RNA, stem cells, and cancer. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **2009**, *87*, 711-746.
3. Sauvageau, M.; Sauvageau, G. Polycomb group genes: keeping stem cell activity in balance. *PLoS biology* **2008**, *6*, e113.
4. Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M., Jr.; Jungkamp, A.C.; Munschauer, M.; Ulrich, A.; Wardle, G.S.; Dewell, S.; Zavolan, M.; Tuschl, T. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **2010**, *141*, 129-141.
5. Yang, J.H.; Li, J.H.; Shao, P.; Zhou, H.; Chen, Y.Q.; Qu, L.H. starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. *Nucleic acids research* **2011**, *39*, D202-209.
6. Gregory, P.A.; Bert, A.G.; Paterson, E.L.; Barry, S.C.; Tsykin, A.; Farshid, G.; Vadas, M.A.; Khew-Goodall, Y.; Goodall, G.J. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nature cell biology* **2008**, *10*, 593-601.
7. Korpal, M.; Lee, E.S.; Hu, G.; Kang, Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *The Journal of biological chemistry* **2008**, *283*, 14910-14914.
8. Gibbons, D.L.; Lin, W.; Creighton, C.J.; Rizvi, Z.H.; Gregory, P.A.; Goodall, G.J.; Thilaganathan, N.; Du, L.; Zhang, Y.; Pertsemlidis, A.; Kurie, J.M. Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. *Genes & development* **2009**, *23*, 2140-2151.
9. Korpal, M.; Ell, B.J.; Buffa, F.M.; Ibrahim, T.; Blanco, M.A.; Celia-Terrassa, T.; Mercatali, L.; Khan, Z.; Goodarzi, H.; Hua, Y.; Wei, Y.; Hu, G.; Garcia, B.A.; Ragoussis, J.; Amadori, D.; Harris, A.L.; Kang, Y. Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. *Nature medicine* **2011**, *17*, 1101-1108.
10. Barroso-delJesus, A.; Romero-Lopez, C.; Lucena-Aguilar, G.; Melen, G.J.; Sanchez, L.; Ligero, G.; Berzal-Herranz, A.; Menendez, P. Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. *Molecular and cellular biology* **2008**, *28*, 6609-6619.
11. Lin, S.L.; Chang, D.C.; Chang-Lin, S.; Lin, C.H.; Wu, D.T.; Chen, D.T.; Ying, S.Y. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA* **2008**, *14*, 2115-2124.
12. Shu J, Xia Z, Li L, Liang ET, Slipek N, Shen D, Foo J, Subramanian S, Steer C. Dose-dependent differential mRNA target selection and regulation by let-7a-7f and miR-17-92 cluster microRNAs. *RNA Biol* ***2012,*** 9:1275-1287.

Appendices: PDF files of an article published and a draft manuscript.

MicroRNA mediated Gene Regulatory Networks in Malignant Peripheral Nerve Sheath Tumors

Kwangmin Choi[1§], Anil G Jegga [2], Praveensingh Hajeri [3], Subbaya Subramanian [3], Nancy Ratner[1]

Abstract

Background

One of the main hurdles to develop effective therapy for malignant peripheral nerve sheath tumors (MPNST) is the lack of understanding of molecular mechanisms regulating cancer genes and metastasis. Oncogenic epithelial-to-mesenchymal transition (EMT) is a critical step for metastasis and is closely related with transcriptional changes of many key genes involved in cell polarity, cell-cell adhesion, and cell migration. Multiple abnormal signaling in this pathways and processes can initiate, promote, and maintain EMT process. To understand the molecular networks that regulates EMT pathway and promote metastasis in MPNST, we systematically analyzed the microRNA (miRNA)–mediated gene regulatory networks using miRNA and MRNA expression profiles generated from normal Schwan cells, MPNSTs tumor tissues and cell lines.

Results

Both negative- and positive- correlation networks based on gene and miRNA expression data were generated. These correlations were further mined by referring to various relational database (protein-protein interactions, canonical pathways, transcription factor-to-target prediction) and genomics data (copy-number alteration, differential DNA methylation). We identified six candidate network modules, which potentially control the preferential activation of TGF-beta/SMAD signaling to TGF-beta/non-SMAD signaling and the induction of cancer cell stemness.

Conclusions

By applying different levels of data integration and exploration, we could identify several units which take part in EMT of MPNST. Reconstructed networks from this study suggests that miRNAs actively participate in transcription control of cancer

genes and cause aberrant modification of core pathways responsible for transformation and metastasis in MPNST development.

Background

Neurofibromatosis type 1 (NF1) caused by germline mutations of the *NF1* gene is an autosomal dominant disorder affecting one in 3000 individuals world-wide. NF1 [1]. Benign tumors neurofibroma (NF) originates from Schwann cells in NF1 patients. About 95% of NF patients have multiple dermal NFs (dNF), 30% develop plexiform NFs (pNF), and 5 -10 % develop malignant peripheral nerve sheath tumors (MPNSTs). MPNST patients' survival rate beyond 5 years is less than 43% [2, 3]. Although it is believed that pNF may transform to MPNST, the sequences of biological events and exact mechanisms driving malignant tumor is not completely understood. Beyond mutations in both copies of the *NF1* gene, few molecular changes have been associated with MPNSTs [4]. One of the main hurdles to develop effective therapy for MPNST is the lack of understanding of deregulation in miRNA mediated molecular networks that lead to uncontrolled cell growth and invasion. distinguishing characteristic of MPNST is its propensity to metastasize to other body sites, predominately to the lungs. Epithelial-to-Mesenchymal Transition (EMT) is a process where fully differentiated and polarized epithelial cells transform their epithelial phenotypes into mesenchymal ones, depending on a set of transcription factors and cellular environment. EMT is generally characterized by loss of cell polarity, loss of epithelial markers (e.g. cell-to-cell adhesion proteins), gain of mesenchymal markers (e.g. α-smooth muscle actin), cytoskeletal remodeling, and, and migratory phenotype [5].

EMT plays a critical roles in both normal cell development programs and cancer metastasis [6]. Oncogenic EMT is a result of various cellular changes

including aberrant apoptosis/anoikis and modified molecular signaling. Among many abnormal signaling pathways, RAS activation [7] and TGF-beta pathway [8] has been known as central players in oncogenic EMT. Autocrine induction of TFG-beta and its interaction with RAS significantly promote EMT and migration of malignant cells. The cancer type-dependent differential roles of SMADs in TGF-beta-induced EMT have been also suggested [8].

MicroRNAs (miRNAs) play a major role in the post-transcriptional regulation of cancer genes [9-13]. Although the biological roles and mechanisms of many miRNAs are still not fully understood, two independent regulatory mechanisms such as RNA inference (RNAi) and RNA activation (RNAa) have been investigated. In RNAi mode [14], a miRNA can cleave target mRNAs (in case of perfect complementary match with its target mRNAs) or inhibit protein translation from the target mRNAs (in case of imperfect complementary match to its target mRNAs) [15]. Whereas in RNAa mode, promoter-targeting miRNAs can induce prolonged activation of gene expression associated with epigenetic changes. Aberrant RNAa in pathological condition has been suggested as another level of epigenetic control mechanism in cancers [16-18]. Actual transcriptional and/or translational regulation by a group of miRNAs can be very complex because multiple miRNAs may cooperatively regulate one target mRNA or a single miRNA multiple target genes.

DNA copy-number alteration and/or epigenetic changes in cancer genomes can also significantly modify gene regulatory programs via miRNAs [16, 17]. Furthermore, it is possible that the same miRNA can switch its role between RNAi and RNAa, depending on a given cellular contexts and target gene availability.

We reconstructed miRNA-gene correlation networks of MPNST transcriptome, using genome-wide gene and miRNA expression data. Our strategy is a similar approach

used in MMIA (need to expand) [19], which is specified to the data to with a control/disease contrast. To focus on cancer genes whose expression are potentailly controlled by miRNAs, we also applied two additional filters based on available DNA copy number aberration and DNA methylation data. Through this genome-wide integrative approach we uncovered several key regulatory units involved in oncogenic EMT pathway in MPNST.

Methods

Gene expression data and DE-genes, (Figure 1-B)

The gene expression data used in this study is from our previous study ([4], GEO accession # GSE14038, whole-genome Affymetrix GeneChip HU133 Plus 2.0 using the Affymetrix protocol). Microarray data from 10 independent primary human Schwann cells (NHSC, n = 10) and MPNST cell line (MPNST, n = 13) from patient samples were used for this study.

Statistical comparisons were done using R/Bioconductor and GeneSpring GX v7.3.1 (Agilent Technologies). Differentially expressed (DE)-genes were defined as genes whose expression levels were at least three-fold higher or lower in target groups (MPNST) compared to NHSC after applying Benjamini and Hochberg false discovery rate [20] correction (FDR/BH $p \leq 0.05$) . For gene annotation, custom CDF (custom GeneChip library file) based on RefSeq target definitions (Hs133P REFSEQ Version 8) was downloaded and used to provide accurate interpretation of GeneChip data [21].

MicroRNA expression array and DE-miRNAs (Figure #2, #2)

Three out of 10 NHSC samples and three out of 13 MPNST samples were used to extract miRNAs for miRNA expression microarray experiment. Total RNAs were extracted from NHSC and MPNST cells using miRvana total RNA isolation kit

(Ambion) according to the manufacture's instruction. We used Illumina Sentrix Array Matrix for miRNA expression profiling as previously described [22]. Briefly, total RNA quality was determined by Agilent 6000 nanochip (Agilent Technologies, Palo Alto, CA, USA). Only samples with RNA index number (RIN) of 6 were subjected to miRNA analysis using miRNA BeadArrays. 500 ng of total RNA was used for each sample. The miRNA BeadArray procedure [22] is similar to the cDNA-mediated annealing, selection, extension and ligation (DASL) method [23]. After hybridization, the arrays were imaged using an Illumina BeadArray Reader and the fluorescent intensity of miRNA probes were analyzed using BeadStudio version 3.1.1(Illumina).MiRNA array was performed using the human MO_V2 chip (Illumina), which contains up to 1145 miRNAs.

Per-chip normalization to 50 percentile and then per-gene median normalization were applied to the raw data on GeneSpring GX v7.3.1. DE-miRNAs were defined as miRNAs whose expression levels were at least 1.5-fold higher or lower in MPNST compared to NHSC, with t-test p-value <0.05 (without FDR correction). When the fold change of a given miRNA was bigger (or smaller) than 1500 (or 1/1500), the DE-miRNA was considered to be artifact and excluded from the analysis.

MiR target prediction (Figure 3, #3)

Once DE-miRNAs were identified, their putative target genes were matched among DE-genes by referring to three miR-target databases. We selected (i) PITA (http://genie.weizmann.ac.il/pubs/mir07/mir07_data.html, 3/15-flank, TOP prediction) [24]; (ii) TargetScan-human (v5.2, conserved targets) [25]; and (iii) miRanda/mirSVR (http://www.microrna.org/microrna, August 2010, conserved and good mirSVR < -0.5) [26] for compiling the putative targets of DE-miRNAs.

Additionally, a miR-to-target gene relationship was considered only if the prediction was found in at least two of three selected databases. Based on retrieved miRNAs and their target genes, initial miR-to-gene correlation networks were generated.

Correlation analysis using Pearson-correlation between paired sample data (Figure 4, #4)

Resulting bipartite graphs/networks (i.e. miR-target networks) are complex and do not necessarily reflect true regulatory networks based on real transcriptome profile of MPNST genome. To reduce network complexity and leave out weak correlation edges (or false positives) in the networks, Pearson correlation-based filtering was applied. For each miR-to-target pair, Pearson correlation coefficient was calculated between three paired samples (i.e. three gene microarray and their paired miRNA microarray data). The correlation between two variables reflects the degree to which the variables are related. Strong ($0.4<|r|\leq 0.7$) and very strong ($0.7<|r|\leq 1.0$) correlations were only considered. Pearson-correlation coefficient was computed following conventional formula.

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{N}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{N}\right)}}$$

Network analysis (Figure 5, #5-6 and Figure 2-4)

A visualization program was written in Python and NetworkX package (http://networkx.lanl.gov, Ver. 1.6) to easily explore data and visualize hidden cluster structures (NEATO layout).

To visualize the coordinated participation of multiple miRNAs on the same biological theme, protein-protein interaction (PPI) and pathway membership information were also integrated. Gene sets of three canonical pathway databases (KEGG (http://www.genome.jp/kegg), REACTOME (http://www.reactome.org),

BIOCARTA (http://www.biocarta.com) were extracted from MSigDB v.3.0 (http://www.broadinstitute.org/gsea/msigdb) and a red edge was assigned to two genes when two genes in a network belong to the same canonical pathway. A blue edge connects two genes if their protein products directly interact in manually curated HPRD database (Release 9; http://www.hprd.org). We considered using only manually curated PPI data set from HPRD to avoid potential false positives. By adding this information, the resulting networks could reveal several clues of how genes and miRNAs are functionally interrelated in the networks. Finally, cancer-related network genes from the most updated cancer-related gene list from Bushman Lab (2032 genes, Release 8/2011; http://microb230.med.upenn.edu/protocols/cancergenes.html) were highlighted with yellow (up-regulated genes) or green (down-regulated genes) colors. The final networks contain the following information: (i) fold change levels of DE-genes and DE-miRNAs (size of node); (ii) strong correlated DE-genes and DE-miRNAs; (iii) pathway memberships of DE-genes; and (iv) physical interaction between DE-genes (see Figure 2-5). The strong negative- and strong positive- correlation networks two types of correlation networks were merged into a single network as in Figure 6. For better visibility, PPI and gene association (canonical pathways) information was removed in Figure 7-A. A subset network in Figure 4 was generated using only cancer genes and their targeting miRNAs from the original combined correlation network.

A network of miRNA-controlled TFs and their target genes (Figure 1-G and Figure 5).

Some of miRNAs in the correlation network seem to participate in the coordinated transcriptional controls on a set of transcription factors in order to maximize downstream effects. From a main correlation network in Figure 4, four differentially

expressed transcription factors (DE-TFs) were identified and a new subset network was constructed using these four TFs, their target genes, and DE-miRNAs showing strong correlation with these genes (Figure 5-A). Figure 5-B is a simplified version of Figure 6-A using several target genes/miRNAs of interest.

Based on 615 MSigDB C3-TFT gene set (originally from Xie et al. [27] and Transfac v7.4 (http://gene-regulation.com)), TFs and their target genes in the combined network were extracted and then a new network was reconstructed. Three edge types in these networks are (i) miR-to-target gene relations (gray color); (ii) TF-vs-target gene relations (orange); (iii) canonical pathway membership (red); and (iv) HPRD PPI (blue). Gene set enrichment test of target genes for each TF was performed using ToppGene Suite (http://toppgene.cchmc.org).

Whole-genome methylome data (see Figure 8-H)

Ferber et al. [28] recently published unbiased whole-methylome data of normal primary human Schwann cell, benign NF and malignant MPNST genomes. Their MeDIP-seq data have been deposited in GEO (http://www.ncbi.nlm.nih.gov/geo) under accession number GSE21714. The method detecting differentially methylated regions (DMR) in MPNST compared to NHSC was adopted from Ferber et al. [28]. Briefly, Batman methylation scores per 100 bp were averaged for each 1K bp windows. A conservative threshold for calling DMR calling was used based on the 95th percentile of the difference in methylation score. DMR regions were mapped to human genome hg18 version (Build NCBI-36).

The nearest CpG island shores (CpG-IS) to the transcription start sites of each gene/miRNA were scanned [28, 29]. The definition of CpG-IS is areas up to 2K bp distant from CpG islands. We consider only the nearest CpG-IS from the transcription

start site (TSS), within 5K bp ranges from each TSS. The genomic coordinates of miRNAs, genes, and CpG islands (NCBI36/hg18) were extracted from corresponding tracks of UCSC Genome Browser (http://genome.ucsc.edu). In case of intragenic miRNAs, we assumed that their expression is influenced by the nearest CpG-IS to the transcription start site of their host gene (see Supplementary Data Table 1)

Copy-number alteration data (see Figure 9-I)

Unpublished copy-number alteration (CNA) data obtained from 14 MPNST patient tumor samples were provided by Dr. Eduard Serra at Institut de Medicina Predictiva i Personalitzada del Càncer (Barcelona, Spain). The CNA regions were mapped to human genome hg18 version (build NCBI-36). Since no statistical scores are assigned to any region yet, we simply assumed that a gene (region) is gained (or lost) when the whole gene coding region falls within any reported gained (or deleted) region. Since this condition is very stringent, any conclusion from CNA data needs to be considered provisional.

The possible relationships of expression fold change, CNA, DNA methylation of all genes and miRNAs in our correlation network are provided as supplementary data (See Supplementary Data Table 1, Excel format). Literature search (Figure 1-J)Many portions of gene-to-gene, miR-to-gene, or protein-to-protein interaction data have not fully mined from biomedical literatures yet. To efficiently search and collect known or inferred relationships between gene-gene, protein-protein, or miR-gene, each edge (e.g., miR-gene or gene-gene) in our combined network was used as a query to BioGraph [30]. A publication list per each query was parsed from the HTML output and additional relational information (e.g. activation, inhibition) from the literatures were added to Figure 6. Several gene nodes (e.g. "(MMP1)") were also added to Figure 5 if they seemed to be important players based on collected literatures.

Genes classified by gene expression fold change, copy-number alteration, and DNA methylation (Figure 6, Table 3)

DE-genes in correlation networks were plotted according to their fold change levels (up/down), DNA methylation patterns (hyper-/hypo-methylation), and CNA patterns (gain/loss). The resulting 3-dimensional plots were presented in Figure 5. Here, "fold change" (Z-axis, vertical) means one gene's differential fold change level compared to NHSC. The +/- symbols represent up/down-regulation of the gene. The CNA (Y-axis) score means the ratio of patient sample numbers showing copy number change (gain or loss) to total 14 samples. The ratio used in X-axis (DM) is the ratio of differentially hyper- or hypo-methylated regions (DMR) covered in the nearest CpG-IS region (2K in length). We only consider cases where a given CpG-IS region is covered over 33% by DMR. If CpG-IS regions are covered by both hyper- and hypo-DMRs, it was not considered to be a differentially methylated CpG-IS.

A gene's expression pattern in microarray can be influenced by any of (i) targeting miRNAs, (ii) DNA methylation (=DM) status in CpG-IS regions, and (iii) copy-number alteration (=CNA) status and these effects can make the interpretation of gene/miRNA expression profile complicated. Thus we primarily focused on genes which show no or minute differential methylation and CNA patterns (e.g. LGI1) and then expand genes of interest as needed. In Figure 6, these primary target genes are aligned vertically (Z-axis, FC) on the center of X-Y (DM-CNA) plane.

Identification of key regulatory units

Several genes of interest were chosen from the correlation network, then functional annotations (e.g. "induces", "represses", "binds", "acetylates", etc.) between two genes were extracted from the published literatures and added to edges in proposed regulatory units (Figure 5-B and Figure 7). Thus, all gene-to-gene edges and some

miR-to-gene edges in these units represent experimental evidences on their functional relations, although two genes (or a miRNA and its target gene) do not necessarily have the sample functional relations in MPNST because most evidences were from non-MPNST cancers. Thus all edge annotation need to be considered to be provisional.

Results and Discussion

DE-miRNAs and DE-genes

47 up-regulated and 91 down-regulated miRNAs were detected by comparing to NHSC (fold change > 1.5 or < 1/1.5, $p<0.05$ without FDR correction). 431 up-regulated and 520 down-regulated were identified by comparing to NHSC (fold change > 3 or <1/3, $p<0.05$ with FDR (=BH) correction).

Before and after applying Pearson correlation filter (Table 1)

By searching TargetScan, PITA, and miRanda/mirSVR databases, miRNA:target relationships between DE-miRNAs and DE-genes were retrieved. Resulting networks have high complexity (see Table 1) and possibly contains many false positives because DB-based target prediction does not necessarily mean the retrieved miR:gene pairs are true regulatory units in a given MPNST transcriptome. Pearson correlation-based filter could significantly reduce network complexity and provide a degree of correlation between a miRNA and its target gene in transcription level. For example, the negative correlation network (down-regulated miRNAs and up-regulated genes) had 44 nodes of miRNAs, 335 nodes of genes, and 963 edges between two node types. After filtering, the same network could be effectively pruned to 12 nodes of miRNAs, 45 nodes of genes, and 56 edges between them (see Table 1).

Correlation networks with PPI and pathway membership information (Figure 2-4 and Table2)

Five up-regulated miRNAs (miR-148a, miR-301a, miR-92b, miR-101, and miR-96) and 11 down-regulated miRNAs (miR-216b, miR-338-3p, miR-365, miR-449b, miR-424, miR-200c, and let-b/d/f, mir-211, miR-365) are involved in both negative and positive correlation networks. Three up-regulated miRNAs (miR-242-3p, miR-335, and miR-551b) were only involved in a positive correlation networks while miR-542-3p belongs only to a negative correlation network. Down-regulated miR-450a belongs to positive correlation network.

Transcription factors and their targets (Figure 5, supplementary Table 3)

ETS2 (4.6x) shows strong negative correlation with miR-200c (0.38x). 13 genes (including ETS2 itself) in the correlation network have a conserved ETS2 binding site within promoter region and 12 genes show significant enrichment result (p-value=2.783E-12) on RYTTCCTG_V$ETS2_B term. SOX9 (47x) is negatively correlated with miR-216b (0.64x) and connects to 11 TF target genes (including SOX9 itself) found the correlation network. 10 out of 11 target gene set show significant enrichment result (p-value =5.505E-15) on CATTGTYY_V$SOX9_B1 term. PITX2 (7.2x) is negatively correlated with miR-211 (0.041x). 25 target genes (including PITX2 itself) in the correlation network show very significant enrichment result (p-value=1.349E-33) on GGATTA_V$PITX2_Q2 term (1.349E-33). Finally, PAX3 (0.32x) is negatively correlated with miR-92b (1.9x) and targeting 6 genes (including PAX3 itself) found in the correlation network. 5 out of 6 target genes also shows significant enrichment result (p-value=1.036E-9) on CGTSACG_V$PAX3_B term.

MPNST, metastasis, and EMT

The activation of EMT pathway triggers invasion and migration of cancer cells. Since the most distinguishing characteristic of MPNST is its aggressive invasion to different body sites, understanding of how MPNST cells modulate EMT/MET pathways can be a critical step to find effective therapeutic targets against this invasive cancer.

We investigated how DE-miRNAs possibly control targeted cancer genes in MPNST and several miRNAs were identified as promising candidates involved in epigenetic modulation of key EMT gene expression. Although positive correlation between miRNAs and target genes may imply RNA activation mechanism, we will mainly focus on negatively correlated cases in Discussion section because it is still difficult to interpret positive correlation as RNA activation yet. For example, when a miRNA and its target gene are both up-regulated, the transcriptional up-regulation of the target gene can be a result of direct RNA activation or an outcome of the indirect mechanism where an unknown suppressor to the target is repressed by the miR. (see Figure 6 and 7).

MiR-200c feedback loop and TGF-beta signaling pathway

Five members of the miR-200 family are encoded in the two gene loci: miR-200b (0.37x)-200a (0.0049x)-429 (0.48x) and miR-200c (0.38x)-141 (0.55x). MiR-141, miR-200b, and miR-200c are highly conserved in sequence level. In our MPNST model, all five miR-200 family genes were down-regulated.

The miR-200 family genes are known to be strong inducers of epithelial differentiation and directly inhibit the expression of ZEB1 and ZEB2 [31, 32]. ZEB1 and ZEB2 repress cell-adhesion and polarity-target genes [33, 34]. Wellner et al. [35] reported that ZEB1 also inhibits the expression of the miR-200 family genes and causes EMT-activation and stemness-maintenance by suppressing stemness-inhibiting miRNAs such as miR-200s.

ZEB1 is one of the central regulator genes which control EMT pathway [36] and acts as a transcriptional repressor or activator depending on several poorly understood conditions. Up-regulated ZEB1 [37] is known to repress CDH1 (a.k.a. E-cadherin expression [38], cell junction, and cell polarity [39] in many types of cancers. Although CDH1 loss is generally linked to increased EMT and metastasis both *in vitro* and *in vivo*, our MPNST data shows no significant change in the expression level of CDH1. It is probably related with the stage of EMT [40] or preferential activation of TGF-beta/SMAD pathway [8]. In our correlation network (Figure 4), ZEB1 (5.5x) shows strong negative correlation with miR-200c (0.38x) and strong positive correlation with miR-342-3p (2.0x). Interestingly, miR-342-3p (2.0x) targets both PDGFRA (51x) and ZEB1 (5.5x). It is not clear whether this observation is a result of RNA activation or not. It is also possible that the expression of unknown suppressor (s) of ZEB1 and PDFGRA is repressed by miR-342-3p. Interestingly, ZEB2 is down-regulated in our MPNST data. There is one report that ZEB1 and ZEB2 undergo opposing roles in TGF-beta/BMP/SMAD pathway, that is, ZEB1 enhances but ZEB2 represses SMADs-mediated gene expression in TGF-beta/BMP/SMAD pathway [41]. In our MPNST data, ZEB1 does not show differential DNA methylation pattern in the nearest CpG-IS region nor copy-number alteration patterns, thus it is plausible that the expression of ZEB1 is mainly controlled by miR-200c (supplementary Table 2).

Many different types of signals can induce EMT, but TGF-beta is considered to be a master switch [8]. TGF-beta seems to be a crucial inducer of ZEB1 expression and EMT progression in MPNST (TGFB1=1.8x, TGFB2=2.5x, TGFB3=3.1x, TGFBR1=2.0x, TGFBR2=1.4x). Gregory et al. [32] proposed that an autocrine TGF-beta singling can induce and maintain EMT process via ZEB/miR-200 loop. It is also

known that over-expression of the miR-200 family alone is sufficient to block TGF-beta-induced EMT [31, 32]. Additionally, Wang et al. [42] reported that all five members of the miR-200 family are significantly down-regulated in cells undergoing EMT in response to TGF-beta signaling. Our MPNST data are well consistent with these reports. Interestingly, two miR-200 family members, miR-141 (0.55x) and miR-200c (0.38x), are known to directly target and negatively regulate TGFB2 (2.5x) gene in cancer cells [43].

Thus, TGFB2, ZEB1, and miR-200c seem to work as a reciprocally regulating unit to initiate and stabilize EMT and to promote invasion of MPNST cancer cells.

MiR-211, miR-503, and miR-338-p: TGF-beta/SMAD3 pathway (Figure 7-B)

The tumor suppressive role of miR-211 in malignant and invasive melanoma has been well studied [44, 45]. In our MPNST data, miR-211 (0.041x) is significantly down-regulated and also negatively correlated with several target cancer genes, which are involved in TGF-beta/SMAD/BMP pathway [46]. In MPNST data, miR-211 is negatively correlated with ZNF423 (6.6x) and ZNF521 (13x). Both ZNF423 and ZNF521 are associates with SMADs in response to BMP2 and can activate the transcription of BMP target genes [47]. In our gene expression data, only SMAD3 (2.8x) is up-regulated, but SMAD2 (0.65x), SMAD4 (0.53x), and SMAD6 (0.83x) are slightly down-regulated. ZNF423 and ZNF521 are known to act as transcriptional activators or repressors, depending on biological context of cells. When ZNF423 and ZNF521 act as repressors to TGF-beta pathway, they interact with EBF1 and repress LTBP1 [47]. Since the expression level of LTBP1 (6.1x) is rather increased, ZNF423 and ZNF521 may not play as repressors of TGF-beta pathway in MPNST. ZNF423 is also negatively correlated with miR-503 (0.3x), which is known to promote cell

differentiation and cell cycle quiescence/G1 arrest and decrease proliferation and metastasis of cancer cells [48].

Mir-211 also shows strong negative correlation with its target PITX2 (7.2x). PITX2 is activated by SMAD2/3/4 via TGF-beta/SMAD pathway, canonical and non-canonical WNT pathways [49, 50]. PITX2 also seems to play a critical role in TF-vs-target network (Figure 7) of MPNST with other central TFs (i.e. SOX9, PAX3, ETS2). The nearest CpG-IS region of PITX2 (7.2x) is hyper-methylated according to GSE21714 data (see Supplementary Table 2). Interestingly, Hirose et al. [51] reported that PITX2 is down-regulated and its expression level shows anti-correlation with metastasis and cell growth of colorectal cancer. Hyper-methylation of PITX2 gene was suggested as a mechanism explaining significant decrease of PITX2 expression in prostate cancer [52]. Interestingly, Gu et al. [53] reported that PITX2 is up-regulated in breast cancer. Thus the effect of hyper-methylation may not fully explain the expression level of PITX2 in MPNST.

MiR-338-3p (0.0094x) is significantly down-regulated in MPNST compared to NHSC. This miRNA is known to be a suppressor of metastasis in liver cancer [54], but its role in MPNST development has not reported yet. According to Gokey *et al.* [55], miR-338 is SOX10-dependent and SOX10 (0.035x) directly regulates the expression of miR-338 by binding to the internal promoter of its host AATK (17x) gene in Schwann cell. Thus this report is consistent with our MPNST data.

EVI1 (7.7x) showed a strong negative correlation with miR-338-3p. EVI1 is an oncogene and increases cell survival by blocking TGF-beta-mediated apoptosis via PI3K/AKT and by interacting with SMAD3 (2.8x) [56]. EV1 is also known to regulate hematopoietic stem cell proliferation [57]. Since the EVI1 gene is gained (2

out of 14 patients, unpublished data) and its promoter regions are hyper-methylated, it is not clear whether miR-338-3p is a single factor controlling up-regulation of EVI1.

Let-7d/f and miR-338-p: Stemness control via EPHA4 (Figure 7-C)

Cancer stem cells (CSC) are believed to have invasive and migratory capacity required for metastasis and several miRNAs have been reported that they can actually regulate metastasis via stemness control [58, 59]. Several well-known pluripotent stem cell genes [11] are up-regulated in our MPNST data: CCND2 (3.72x, negatively correlated with miR-503), MYC (2.6x), LIN28B (8.1x), and HMGA2(44.17x) .

Let-7 members are good candidates showing the strong relationship between stemness and metastasis. For example, let-7 family can inhibit self-renewal and maintenance capacity in undifferentiated status of breast cancer cells [11]. In our MPNST data, two let-7 family miRNAs (let-7d/f) seem to be involved in one type of stemness control along with miR-200c/ZEB1/TGFB2 module. Let-7b (0.36x) and let-7f (0.59x) show both positive- and negative- correlations with many cancer-related genes (see Figure 4). Both let-7b and let-7f show strong negative correlation with EPHA4 (3.0x), which is a potential regulator of neurogenesis and causes a loss of cell polarity via abnormal WNT/PCP pathway [60]. Sperber et al. [61] reported that EPHA4 is expressed only by neural stem cells (NSCs) in adult neurogenic niches and directly binds to FYN (0.29x) and causes undifferentiated and unmyelinated axons [62]. EFNB2 (3.2x), showing strong positive correlation with miR-148a (2.2x), directly binds to EPHAs with high affinity [63]. In addition, FGFR1 (3.2x), which is negatively correlated with miR-424 (0.54x), directly interacts with EPHA4 [64] and initiates EMT in prostate cancers [65, 66]. Thus EPHA4 seems to be a key player in stemness control during EMT in MPNST.

Non-canonical WNT signals are transduced to the WNT/planar cell polarity (PCP) signaling pathway, and aberrant WNT/PCP pathway can cause tumor metastasis [67] because WNT/PCP pathway controls tissue polarity, cell adhesion, and motility of developing cells. Thus aberrant changes in WNT/PCP pathway may lead to more severe malignant phenotypes including severe MPNST metastasis. CELSR2 (0.22x) , a member of atypical cadherin family, and FZD3 (0.32x), a representative WNT receptor in WNT/PCP pathway, are known to be key players in neuronal migration [68]. These two WNT/PCP pathway genes show strong positive correlations with miR-338-3p. There is no report yet if miR-338-3p can regulate target genes in RNAa mode.

MiR-92b and miR-196: activation of PDGFRA (Figure 7-D)

In humans, two polycistronic miR-17-92 cluster [69] and miR-106A-363 clusters [70] encode two different miR-92 loci (miR-92a in chromosome 13 and miR-92b in chromosome 1), respectively [71-73]. While the oncogenic role of miR-92a (not differentially expressed) in miR-17-92 cluster (chromosome 13) has been well studied [74], the role of miR-92b (1.9x) is still unclear. Since miR-92a and miR-92b differ by only one nucleotide within their seed sequences, it is possible that miR-92b, instead of miR-92a, may play an oncogenic role in MPNST development. Interestingly, miR-92b was reported to be over-expressed in neuronal precursors and stem cells compared to adult brain [75].

Pavan et al. [76] reported that SATB1 regulates gene expression by acting as a "docking site" for chromatin remodeling enzymes and also by recruiting co-regulators directly to promoters. When SATB1 interacts with KAT2B (a.k.a. PCAF, no copy-number alteration, no differential DNA methylation), KAT2B acetylates the PDZ-like domain of SATB1 and this leads to the loss of its DNA binding activity, and

finally leads to the down-regulation of genes in the vicinity. KAT2B also acetylates the histone proteins, so directly regulates the gene expression. In addition, two protein-level regulations by SATB1 and TWIST1 (36x) are also identified. This suggests that a strong driving force acts on KAT2B during in EMT of MPNST.

In our MPNST data, the expression levels of both SATB1 (0.2x) and KAT2B (0.1x) are significantly down-regulated, so the transcription of some set of genes (possibly, signature genes with epithelial features) in the target regions can be probably down-regulated. In our network, three independent miRNAs target KAT2B, including miR-92b (1.9x, negative correlation), miR-200c (0.38x, positive correlation), and miR-342-3p (2.0x, negative correlation).

Grady et al. [77] reported that miR-342 (2.0x), which is encoded in an intron of the gene EVL, is commonly suppressed in human colorectal cancer. Wang et al. [78] also reported that miR-342 inhibits proliferation and metastasis of colorectal cancer cells by directly targeting DNA methyltransferase 1. In contrast to their reports, our MPNST data shows that both miR-342 (2.0x) and its host gene EVL (2.0x) show two-fold increases.

MiR-92b (1.97x) shows a strong positive correlation with TWIST1 (36x) in our correlation network. Hamamori et al. [79] reported that TWIST1 binds to CBP/p300 and KAT2B (0.1x), directly regulates its HAT activities, and results in the inhibition of their acetyltransferase activities. Interestingly, our data shows that the expression fold change of CDH1 is not significant between NHSC and MPNST. As described previously, the expression levels of both KAT2B (0.1x) and its interacting protein SATB1 (0.2x) are down-regulated. Additionally, Shiota *et al.* [80] reported that KAT2B can affect oncogenic properties of TWIST1 including EMT, cell growth, metastasis, sensitivity to anti-cancer drugs via acetylating TWIST1. They reported

that disruption of TWIST1 acetylation inhibits YB1 transcription and nuclear localization of TWIST1. According to the temporal/spatial cooperation of TWIST1 and SNAI1 by Tran *et al.* [40], TWIST1, present throughout the primary tumor, is transiently repressed by SNAI1 (0.60x) during EMT initiation stage, then is increased again in later stages of EMT. According to the ratio of TWIST1 (36x) and SNAI1 (0.60x), our MPNST data probably represent the later EMT stage.

PDGFRA (51x) is one of the most interesting cancer genes, possibly positively controlled by miR-196. This gene encodes a cell surface tyrosine kinase receptor for PDGF family. PDGFs are mitogens for cells with mesenchymal origin and function as cell survival factors during EMT. Jechlinger et al. [81] reported that metastatic potential of oncogenic mammary epithelial cells requires an autocrine PDGF/PDGFR loop, which is a result of TGF-beta-induced EMT. Eckert et al. [82] also reported that, in breast cancer, TWIST1 (36x) directly induces transcription of PDGFA. In our MPNST data, TWIST1 shows strong positive correlation with miR-92b (1.9x). The role of TWIST1 during EMT will be described in the following section.

MiR-96 and miR-196: ECM degradation (Figure 7-E)

MiR-96 (410x) is highly up-regulated in our MPNST data and both negatively- and positively- correlated with several cancer genes in our correlation network. Recent reports on miR-96 showed that the expression of miR-96 is up-regulated in non-small cell lung cancer [36] and prostate cancer [83], and plays an critical role in cancer development.

LGI1 (0.018x) is negatively correlated with miR-96 and significantly down-regulated in MPNST data. LGI1 plays a role in suppressing the production of MMP1 and MMP3 via the PI3K/ERK pathway [84] and MMPs are key players in ECM degradation during EMT. In MPNST, MMP1 (22x) and MMP3 (3.1x) are up-

regulated, plausibly due to the down-regulation of LGI1. LGI1 does not show copy-number alteration or differential methylation pattern (Figure 5 and supplementary Table 2), so its transcription seem to be mainly controlled by miR-96.

PRDM16, a transcription factor, interacts with SMAD2/3 and functions as a repressor of TGF-beta signaling by stabilizing the inactive SMAD3-SKI complex on the promoter of TGF-beta target genes [85]. In MPNST, PRDM16 (0.15x) is significantly down-regulated and negatively correlated with miR-96. Since PRDM16 functions as a repressor of TGF-beta signaling, down-regulation of PRDM16 can induce activation of TGF-beta signaling-induced EMT pathway.

The reduction of FYN expression (0.25x) is interesting because FYN is generally required for enhanced metastasis in many cancer types. FYN is over-expressed in multiple human cancers (e.g. prostate, melanoma, pancreatic, glioma, chronic myeloma). Yadav et al. [86] reported that the activation of AKTs was necessary and sufficient for FYN induction by HRAS in various metastatic cancer types. But in MPNST data, FYN is rather down-regulated by four-fold, but expression fold changes of HRAS (1.2x), AKT1 (0.76x) and AKT2 (1.1x) are not significant. Thus, FYN seems to play a different biological role in MPNST metastasis. FYN seems to show gene loss pattern in MPNST (3/14 patients).

The role of miR-196 (5.3x) is still unclear. Although Li et al. [26] suggested a miR-196 as a possible tumor-suppressor due to its strong correlation with HOXC8, they also reported that the miR-196 family is not correlated with cell motility or metastasis status. Interestingly, in the context of glioblastoma multiforme (GBM), Lakomy et al. [87] reported that the expression miR-196 is up-regulated compare to non-tumor brain tissue and only miR-196b (not miR-196a) is positively correlated with overall survival. Guan et al. [88] also reported that miR-196 is up-regulated in

glioblastoma but not in anaplastic astrocytoma plus normal brains. Thus they suggested that miR-196 may play a role in the malignant progression of gliomas. In our data, miR-196 (5.3x) is both positively and negatively correlation with several cancer genes. MiR-196 is positively correlated with several target genes in our network. Guarino et al. [89] reported that transcription factor GATA6 (20x) induces MMP1 (22x) to promote EMT. As mentioned previously, MMP1 is up-regulated by down-regulation of LGI1, possibly by miR-96. LGI1 plays a role in suppressing the production of MMP1 and MMP3 via the PI3K/ERK pathway and our data shows that both LGI1 and PI3K/ERK pathway are down-regulated, but GATA6 is up-regulated. Invadopodia is a cellular structure related with ECM degradation. According to Eckert et al. [82], PDGFRA (51x), positively correlated with miR-196 (5.3x), increases Src kinase activity (without significant changes in transcription level) which leads to invadopodia formation and/or stabilization by phosphorylating invadopodia components by activated Src kinase. Cell migration (Figure 7-F)

ROBO1 is a receptor for SLIT1/2 and a molecular guidance cue in cellular migration [90] and ROBO1 (3.3x) is positively correlated with miR-92b (1.9x) in our MPNST data. SRC activates ABL to stabilize ROBO1 in order to promote cell migration [91]. ROBO1 is significantly up-regulated during EMT of head and neck squamous cell carcinoma (HNSCC) [92] and shows strong correlation with HS6ST2 (8.4x, no copy-number alteration, no differential DNA methylation), which is also responsible for cell migration, in two models of WNT-induced cancers [93]. Interestingly, Fuxe et al. [94] reported that 6-O-sulfated heparan sulfate is required for the activation of SLIT-ROBO signaling. In MPNST, HS6ST2 (8.5x) is up-regulated and negatively correlate with miR-503 (0.30x). NTN4 (3.9x), which is positively correlated with miR-96, is also responsible for axon guidance and cell migration [95].

RDX (0.23x) shows a strong negative correlation with miR-196. Since RDX encodes a cytoskeletal protein linking actin to the plasma membrane, its down-regulation is possibly related with EMT-related structural modification of cancer cells. RDX is also down-regulated in invasive adenocarcinoma [96].

Conclusions

Most of miR-related researches have been focused on miRNAs' transcriptional inhibitory roles on target genes. This correlation may be computationally detected by looking negative correlation of transcription level of a miRNA and its target genes, but its positive correlations with target genes have been often ignored or underestimated partly because only a few RNA activation cases have ever been reported. In this paper, we do not claim that identified positive correlations explain RNAa mechanism because unknown/undetected suppressors of miRNA target genes can be indirectly involved in this regulatory mechanism.

We integrated positive and negative correlations altogether and reconstructed more comprehensive regulatory networks of MPNST transcriptome. By combining microarray data, protein-protein interaction DB, canonical pathway information, TF-target relation, and experimental evidences from literature, we could identify several units which take part in EMT of MPNST. Roughly two regulatory themes/units were revealed from our MPNST data: (1) preferential activation of TGF-beta/SMAD signaling to TGF-beta/non-SMAD signaling and (2) induction of cancer cell stemness.

Willis et al. [8] summarized TGF-beta/SMAD and TGF-beta/non-SMAD signaling pathway involved in EMT. Interestingly, our analysis shows that MPNST regulates EMT preferentially via TGF-beta/SMAD pathway using a set of miRNAs and interaction of target genes, thus TGF-beta/SMAD signaling [8, 97] seems to be key pathway involved in MPNST metastasis. Interestingly, most of key genes in TGF-

beta/non-SMAD-mediated pathway (e.g. CDH1/E-cadherin change), PIK3s/PI3K, AKT1/3, RHOA, PARD6, SNAI1, MAPKs) show no differential expression pattern or are slightly down-regulated. Huge down-regulation of LGI1 (0.018x) by miR-96 (410x) is probably related with down-regulation of PI3K/MAPK pathway. Probably, this observation explains why our data does not show the significant down-regulation of CDH1/E-cadherin although this pattern is generally considered as EMT markers. TGF-beta/SMAD signaling can cause quite wide range of effect because SMADs with low DNA binding affinity require other cofactors with high affinity/specificity for target genes [98]. EPHA4, FGFR1, let-7d/f, and TGFB2/ZEB1/miR-200c feedback loop seem to be main players in stemness induction. Transcriptional crosstalk between TGF-beta/SMAD pathway and stem cell pathway in cancer EMT are also possible via orchestration multiple miRNAs and their target genes.

Reconstructed networks from this study suggests that miRNAs are actively involved in transcription control of cancer genes and cause aberrant modification of core pathways in transformation and metastasis, although our result is provisional until biologists prove underlying regulatory units by experiments.

Authors' contributions

PH and SS generated miRNA microarray raw data. KC, AJ, and NR discussed and designed the data analysis workflow. KC wrote required programs according to this data analysis workflow. KC, AJ, SS, and NR wrote manuscript.

Acknowledgements

The study is in part funded by the grant funding from Department of Defense (W81XWH-10-1-0556) to SS.

References

1. Levy P, Vidaud D, Leroy K, Laurendeau I, Wechsler J, Bolasco G, Parfait B, Wolkenstein P, Vidaud M, Bieche I: **Molecular profiling of malignant peripheral nerve sheath tumors associated with neurofibromatosis type 1, based on large-scale real-time RT-PCR**. *Molecular Cancer* 2004, **3**(1):20.
2. Evans DGR, Baser ME, McGaughran J, Sharif S, Howard E, Moran A: **Malignant peripheral nerve sheath tumours in neurofibromatosis 1**. *Journal of Medical Genetics* 2002, **39**(5):311-314.
3. Zhou H, Coffin CM, Perkins SL, Tripp SR, Liew M, Viskochil DH: **Malignant peripheral nerve sheath tumor - A comparison of grade, immunophenotype, and cell cycle/growth activation marker expression in sporadic and neurofibromatosis 1-related lesions**. *American Journal of Surgical Pathology* 2003, **27**(10):1337-1345.
4. Miller SJ, Jessen WJ, Mehta T, Hardiman A, Sites E, Kaiser S, Jegga AG, Li H, Upadhyaya M, Giovannini M *et al*: **Integrative genomic analyses of neurofibromatosis tumours identify SOX9 as a biomarker and survival gene**. *Embo Molecular Medicine* 2009, **1**(4):236-248.
5. Zavadil J, Bottinger EP: **TGF-beta and epithelial-to-mesenchymal transitions**. *Oncogene* 2005, **24**(37):5764-5774.
6. Thiery JP: **Epithelial-mesenchymal transitions in tumour progression**. *Nat Rev Cancer* 2002, **2**(6):442-454.
7. Janda E, Lehmann K, Killisch I, Jechlinger M, Herzig M, Downward J, Beug H, Grünert S: **Ras and TGFβ cooperatively regulate epithelial cell plasticity and metastasis**. *The Journal of Cell Biology* 2002, **156**(2):299-314.
8. Willis BC, Borok Z: **TGF-β-induced EMT: mechanisms and implications for fibrotic lung disease**. *American Journal of Physiology - Lung Cellular and Molecular Physiology* 2007, **293**(3):L525-L534.
9. He L, Hannon GJ: **MicroRNAs: small RNAs with a big role in gene regulation**. *Nat Rev Genet* 2004, **5**(7):522-531.
10. Kim VN: **MicroRNA biogenesis: coordinated cropping and dicing**. *Nat Rev Mol Cell Biol* 2005, **6**(5):376-385.
11. Esquela-Kerscher A, Slack FJ: **Oncomirs [mdash] microRNAs with a role in cancer**. *Nat Rev Cancer* 2006, **6**(4):259-269.
12. Calin GA, Croce CM: **MicroRNA signatures in human cancers**. *Nat Rev Cancer* 2006, **6**(11):857-866.
13. Chen K, Rajewsky N: **The evolution of gene regulation by transcription factors and microRNAs**. *Nat Rev Genet* 2007, **8**(2):93-103.
14. Lee RC, Feinbaum RL, Ambros V: **The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14**. *Cell* 2004, **S116**(2):843-854.
15. Williams A: **Functional aspects of animal microRNAs**. *Cellular and Molecular Life Sciences* 2008, **65**(4):545-562.
16. Place RF, Li L-C, Pookot D, Noonan EJ, Dahiya R: **MicroRNA-373 induces expression of genes with complementary promoter sequences**. *Proceedings of the National Academy of Sciences* 2008, **105**(5):1608-1613.
17. Huang V, Place RF, Portnoy V, Wang J, Qi Z, Jia Z, Yu A, Shuman M, Yu J, Li L-C: **Upregulation of Cyclin B1 by miRNA and its implications in cancer**. *Nucleic Acids Research* 2011.
18. Chu Y, Yue X, Younger ST, Janowski BA, Corey DR: **Involvement of argonaute proteins in gene silencing and activation by RNAs**

complementary to a non-coding transcript at the progesterone receptor promoter. *Nucleic Acids Research* 2010, **38**(21):7736-7748.

19. Nam S, Li M, Choi K, Balch C, Kim S, Nephew KP: **MicroRNA and mRNA integrated analysis (MMIA): a web tool for examining biological functions of microRNA expression**. *Nucleic Acids Research* 2009, **37**(suppl 2):W356-W362.
20. Benjamini Y, Hochberg Y: **Limma: Controlling the false discovery rate: a practical and powerful approach to multiple testing**. *J R Stat Soc [Ser B]* 1995, **57**:289 - 300.
21. Dai MH, Wang PL, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H *et al*: **Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data**. *Nucleic Acids Research* 2005, **33**(20).
22. Chen J, Lozach J, Garcia EW, Barnes B, Luo S, Mikoulitch I, Zhou L, Schroth G, Fan J-B: **Highly sensitive and specific microRNA expression profiling using BeadArray technology**. *Nucleic Acids Research* 2008, **36**(14):e87.
23. Fan JB, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, Zhang BH, Shen R *et al*: **A versatile assay for high-throughput gene expression profiling on universal array matrices**. *Genome Research* 2004, **14**(5):878-885.
24. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E: **The role of site accessibility in microRNA target recognition**. *Nat Genet* 2007, **39**(10):1278-1284.
25. Grimson A, Farh KK-H, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP: **MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing**. *Molecular Cell* 2007, **27**(1):91-105.
26. Betel D, Koppal A, Agius P, Sander C, Leslie C: **Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites**. *Genome Biology* 2010, **11**(8):R90.
27. Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M: **Systematic discovery of regulatory motifs in human promoters and 3[prime] UTRs by comparison of several mammals**. *Nature* 2005, **434**(7031):338-345.
28. Feber A, Wilson GA, Zhang L, Presneau N, Idowu B, Down TA, Rakyan VK, Noon LA, Lloyd AC, Stupka E *et al*: **Comparative methylome analysis of benign and malignant peripheral nerve sheath tumors**. *Genome Research* 2011, **21**(4):515-524.
29. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M *et al*: **The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores**. *Nat Genet* 2009, **41**(2):178-186.
30. Liekens A, De Knijf J, Daelemans W, Goethals B, De Rijk P, Del-Favero J: **BioGraph: unsupervised biomedical knowledge discovery via automated hypothesis generation**. *Genome Biology* 2011, **12**(6):R57.
31. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y, Goodall GJ: **The mir-200 family and mir-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1**. *Nat Cell Biol* 2008, **10**(5):593-601.
32. Gregory PA, Bracken CP, Smith E, Bert AG, Wright JA, Roslan S, Morris M, Wyatt L, Farshid G, Lim Y-Y *et al*: **An autocrine TGF-β/ZEB/miR-200**

signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. *Molecular Biology of the Cell* 2011, **22**(10):1686-1698.

33. Korpal M, Lee ES, Hu G, Kang Y: **The miR-200 Family Inhibits Epithelial-Mesenchymal Transition and Cancer Cell Migration by Direct Targeting of E-cadherin Transcriptional Repressors ZEB1 and ZEB2**. *Journal of Biological Chemistry* 2008, **283**(22):14910-14914.
34. Davalos V, Moutinho C, Villanueva A, Boque R, Silva P, Carneiro F, Esteller M: **Dynamic epigenetic regulation of the microRNA-200 family mediates epithelial and mesenchymal transitions in human tumorigenesis**. *Oncogene* 2011.
35. Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, Waldvogel B, Vannier C, Darling D, zur Hausen A *et al*: **The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs**. *Nat Cell Biol* 2009, **11**(12):1487-U1236.
36. Lorenzatti G, Huang W, Pal A, Cabanillas AM, Kleer CG: **CCN6 (WISP3) decreases ZEB1-mediated EMT and invasion by attenuation of IGF-1 receptor signaling in breast cancer**. *Journal of Cell Science* 2011, **124**(10):1752-1758.
37. Graham TR, Zhau HE, Odero-Marah VA, Osunkoya AO, Kimbro KS, Tighiouart M, Liu T, Simons JW, O'Regan RM: **Insulin-like Growth Factor-I–Dependent Up-regulation of ZEB1 Drives Epithelial-to-Mesenchymal Transition in Human Prostate Cancer Cells**. *Cancer Research* 2008, **68**(7):2479-2488.
38. Eger A, Aigner K, Sonderegger S, Dampier B, Oehler S, Schreiber M, Berx G, Cano A, Beug H, Foisner R: **DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells**. *Oncogene* 2005, **24**(14):2375-2385.
39. Aigner K, Dampier B, Descovich L, Mikula M, Sultan A, Schreiber M, Mikulits W, Brabletz T, Strand D, Obrist P *et al*: **The transcription factor ZEB1 (delta EF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity**. *Oncogene* 2007, **26**(49):6979-6988.
40. Tran DD, Corsa CAS, Biswas H, Aft RL, Longmore GD: **Temporal and Spatial Cooperation of Snail1 and Twist1 during Epithelial–Mesenchymal Transition Predicts for Human Breast Cancer Recurrence**. *Molecular Cancer Research* 2011, **9**(12):1644-1657.
41. Postigo AA: **Opposing functions of ZEB proteins in the regulation of the TGF beta/BMP signaling pathway**. *Embo Journal* 2003, **22**(10):2443-2452.
42. Wang L, Wang J: **MicroRNA-mediated breast cancer metastasis: from primary site to distant organs**. *Oncogene* 2011.
43. Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, Spaderna S, Brabletz T: **A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells**. *EMBO Rep* 2008, **9**(6):582-589.
44. Levy C, Khaled M, Iliopoulos D, Janas MM, Schubert S, Pinner S, Chen P-H, Li S, Fletcher AL, Yokoyama S *et al*: **Intronic miR-211 Assumes the Tumor Suppressive Function of Its Host Gene in Melanoma**. *Molecular Cell* 2010, **40**(5):841-849.

45. Mazar J, DeYoung K, Khaitan D, Meister E, Almodovar A, Goydos J, Ray A, Perera RJ: **The Regulation of miRNA-211 Expression and Its Role in Melanoma Cell Invasiveness**. *PLoS ONE* 2010, **5**(11):e13779.
46. Padua D, Massague J: **Roles of TGF[beta] in metastasis**. *Cell Res* 2009, **19**(1):89-102.
47. Hata A, Seoane J, Lagna G, Montalvo E, Hemmati-Brivanlou A, Massagué J: **OAZ Uses Distinct DNA- and Protein-Binding Zinc Fingers in Separate BMP-Smad and Olf Signaling Pathways**. *Cell* 2000, **100**(2):229-240.
48. Sarkar S, Dey BK, Dutta A: **MiR-322/424 and -503 Are Induced during Muscle Differentiation and Promote Cell Cycle Quiescence and Differentiation by Down-Regulation of Cdc25A**. *Molecular Biology of the Cell* 2010, **21**(13):2138-2149.
49. Zhou W, Lin L, Majumdar A, Li X, Zhang X, Liu W, Etheridge L, Shi Y, Martin J, Van de Ven W *et al*: **Modulation of morphogenesis by noncanonical Wnt signaling requires ATF/CREB family-mediated transcriptional activation of TGF[beta]2**. *Nat Genet* 2007, **39**(10):1225-1234.
50. Kioussi C, Briata P, Baek SH, Rose DW, Hamblet NS, Herman T, Ohgi KA, Lin C, Gleiberman A, Wang J *et al*: **Identification of a Wnt/Dvl/ β -Catenin $\rightarrow$ Pitx2 Pathway Mediating Cell-Type-Specific Proliferation during Development**. *Cell* 2002, **111**(5):673-685.
51. Hirose H, Ishii H, Mimori K, Tanaka F, Takemasa I, Mizushima T, Ikeda M, Yamamoto H, Sekimoto M, Doki Y *et al*: **The Significance of PITX2 Overexpression in Human Colorectal Cancer**. *Annals of Surgical Oncology* 2011, **18**(10):3005-3012.
52. Vinarskaja A, Schulz WA, Ingenwerth M, Hader C, Arsov C: **Association of PITX2 mRNA down-regulation in prostate cancer with promoter hypermethylation and poor prognosis**. *Urologic Oncology: Seminars and Original Investigations* (0).
53. Gu F, Hsu H-K, Hsu P-Y, Wu J, Ma Y, Parvin J, Huang T, Jin V: **Inference of hierarchical regulatory network of estrogen-dependent breast cancer through ChIP-based data**. *BMC Systems Biology* 2010, **4**(1):170.
54. Huang X-H, Chen J-S, Wang Q, Chen X-L, Wen L, Chen L-Z, Bi J, Zhang L-J, Su Q, Zeng W-T: **miR-338-3p suppresses invasion of liver cancer cell by targeting smoothened**. *Journal of Pathology* 2011, **225**(3):463-472.
55. Gokey NG, Srinivasan R, Lopez-Anido C, Krueger C, Svaren J: **Developmental Regulation of MicroRNA Expression in Schwann Cells**. *Molecular and Cellular Biology* 2012, **32**(2):558-568.
56. Liu Y, Chen L, Ko TC, Fields AP, Thompson EA: **Evi1 is a survival factor which conveys resistance to both TGF[beta]- and taxol-mediated cell death via PI3K//AKT**. *Oncogene* 2006, **25**(25):3565-3575.
57. Yuasa H, Oike Y, Iwama A, Nishikata I, Sugiyama D, Perkins A, Mucenski ML, Suda T, Morishita K: **Oncogenic transcription factor Evi1 regulates hematopoietic stem cell proliferation through GATA-2 expression**. *EMBO J* 2005, **24**(11):1976-1987.
58. El-Haibi CP, Karnoub AE: **Mesenchymal Stem Cells in the Pathogenesis and Therapy of Breast Cancer**. *Journal of Mammary Gland Biology and Neoplasia* 2010, **15**(4):399-409.
59. Li F, Tiede B, Massague J, Kang Y: **Beyond tumorigenesis: cancer stem cells in metastasis**. *Cell Research* 2007, **17**(1):3-14.

60. Winning RS, Wyman TL, Walker GK: **EphA4 Activity Causes Cell Shape Change and a Loss of Cell Polarity in Xenopus laevis embryos**. *Differentiation* 2001, **68**(2–3):126-132.
61. Sperber BR, Boyle-Walsh EA, Engleka MJ, Gadue P, Peterson AC, Stein PL, Scherer SS, McMorris FA: **A unique role for fyn in CNS myelination**. *Journal of Neuroscience* 2001, **21**(6):2039-2047.
62. Khodosevich K, Watanabe Y, Monyer H: **EphA4 preserves postnatal and adult neural stem cells in an undifferentiated state in vivo**. *Journal of Cell Science* 2011, **124**(8):1268-1279.
63. Bowden TA, Aricescu AR, Nettleship JE, Siebold C, Rahman-Huq N, Owens RJ, Stuart DI, Jones EY: **Structural Plasticity of Eph Receptor A4 Facilitates Cross-Class Ephrin Signaling**. *Structure* 2009, **17**(10):1386-1397.
64. Yokote H, Fujita K, Jing X, Sawada T, Liang S, Yao L, Yan X, Zhang Y, Schlessinger J, Sakaguchi K: **Trans-activation of EphA4 and FGF receptors mediated by direct interactions between their cytoplasmic domains**. *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(52):18866-18871.
65. Acevedo VD, Gangula RD, Freeman KW, Li R, Zhang Y, Wang F, Ayala GE, Peterson LE, Ittmann M, Spencer DM: **Inducible FGFR-1 Activation Leads to Irreversible Prostate Adenocarcinoma and an Epithelial-to-Mesenchymal Transition**. *Cancer Cell* 2007, **12**(6):559-571.
66. Ahmad I, Iwata T, Leung HY: **Mechanisms of FGFR-mediated carcinogenesis**. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* (0).
67. Wang Y: **Wnt/Planar cell polarity signaling: A new paradigm for cancer therapy**. *Molecular Cancer Therapeutics* 2009, **8**(8):2103-2109.
68. Wada H, Tanaka H, Nakayama S, Iwasaki M, Okamoto H: **Frizzled3a and Celsr2 function in the neuroepithelium to regulate migration of facial motor neurons in the developing zebrafish hindbrain**. *Development* 2006, **133**(23):4749-4759.
69. Olive V, Jiang I, He L: **mir-17-92, a cluster of miRNAs in the midst of the cancer network**. *International Journal of Biochemistry & Cell Biology* 2010, **42**(8):1348-1354.
70. Landais S, Landry S, Legault P, Rassart E: **Oncogenic potential of the miR-106-363 cluster and its implication in human T-cell leukemia**. *Cancer Research* 2007, **67**(12):5699-5707.
71. Petrocca F, Vecchione A, Croce CM: **Emerging Role of miR-106b-25/miR-17-92 Clusters in the Control of Transforming Growth Factor β Signaling**. *Cancer Research* 2008, **68**(20):8191-8194.
72. Mendell JT: **miRiad roles for the miR-17-92 cluster in development and disease**. *Cell* 2008, **133**(2):217-222.
73. Li N, Wei C, Olena AF, Patton JG: **Regulation of endoderm formation and left-right asymmetry by miR-92 during early zebrafish development**. *Development* 2011, **138**(9):1817-1826.
74. Croce CM: **Causes and consequences of microRNA dysregulation in cancer**. *Nat Rev Genet* 2009, **10**(10):704-714.
75. Kapsimali M, Kloosterman W, de Bruijn E, Rosa F, Plasterk R, Wilson S: **MicroRNAs show a wide diversity of expression profiles in the developing and mature central nervous system**. *Genome Biology* 2007, **8**(8):R173.

76. Pavan KP, Purbey PK, Sinha CK, Notani D, Limaye A, Jayani RS, Galande S: **Phosphorylation of SATB1, a global gene regulator, acts as a molecular switch regulating its transcriptional activity in vivo**. *Molecular Cell* 2006, **22**(2):231-243.
77. Grady WM, Parkin RK, Mitchell PS, Lee JH, Kim YH, Tsuchiya KD, Washington MK, Paraskeva C, Willson JKV, Kaz AM *et al*: **Epigenetic silencing of the intronic microRNA hsa-miR-342 and its host gene EVL in colorectal cancer**. *Oncogene* 2008, **27**(27):3880-3888.
78. Wang H, Wu J, Meng X, Ying X, Zuo Y, Liu R, Pan Z, Kang T, Huang W: **MicroRNA-342 inhibits colorectal cancer cell proliferation and invasion by directly targeting DNA methyltransferase 1**. *Carcinogenesis* 2011, **32**(7):1033-1042.
79. Hamamori Y, Sartorelli V, Ogryzko V, Puri PL, Wu HY, Wang JYJ, Nakatani Y, Kedes L: **Regulation of histone acetyltransferases p300 and PCAF by the bHLH protein twist and adenoviral oncoprotein E1A**. *Cell* 1999, **96**(3):405-413.
80. Shiota M, Yokomizo A, Tada Y, Uchiumi T, Inokuchi J, Tatsugami K, Kuroiwa K, Yamamoto K, Seki N, Naito S: **P300/CBP-associated factor regulates Y-box binding protein-1 expression and promotes cancer cell growth, cancer invasion and drug resistance**. *Cancer Science* 2010, **101**(8):1797-1806.
81. Jechlinger M, Sommer A, Moriggl R, Seither P, Kraut N, Capodiecci P, Donovan M, Cordon-Cardo C, Beug H, Grunert S: **Autocrine PDGFR signaling promotes mammary cancer metastasis**. *Journal of Clinical Investigation* 2006, **116**(6):1561-1570.
82. Eckert MA, Yang J: **Targeting invadopodia to block breast cancer metastasis**. *Oncotarget* 2011, **2**(7):562-568.
83. Schaefer A, Jung M, Mollenkopf H-J, Wagner I, Stephan C, Jentzmik F, Miller K, Lein M, Kristiansen G, Jung K: **Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma**. *International Journal of Cancer* 2010, **126**(5):1166-1176.
84. Kunapuli P, Kasyapa CS, Hawthorn L, Cowell JK: **LGI1, a Putative Tumor Metastasis Suppressor Gene, Controls in Vitro Invasiveness and Expression of Matrix Metalloproteinases in Glioma Cells through the ERK1/2 Pathway**. *Journal of Biological Chemistry* 2004, **279**(22):23151-23157.
85. Takahata M, Inoue Y, Tsuda H, Imoto I, Koinuma D, Hayashi M, Ichikura T, Yamori T, Nagasaki K, Yoshida M *et al*: **SKI and MEL1 Cooperate to Inhibit Transforming Growth Factor-β Signal in Gastric Cancer Cells**. *Journal of Biological Chemistry* 2009, **284**(5):3334-3344.
86. Yadav V, Denning MF: **Fyn Is Induced by Ras/PI3K/Akt Signaling and Is Required for Enhanced Invasion/Migration**. *Molecular Carcinogenesis* 2011, **50**(5):346-352.
87. Lakomy R, Sana J, Hankeova S, Fadrus P, Kren L, Lzicarova E, Svoboda M, Dolezelova H, Smrcka M, Vyzula R *et al*: **MiR-195, miR-196b, miR-181c, miR-21 expression levels and O-6-methylguanine-DNA methyltransferase methylation status are associated with clinical outcome in glioblastoma patients**. *Cancer Science* 2011, **102**(12):2186-2190.
88. Guan Y, Mizoguchi M, Yoshimoto K, Hata N, Shono T, Suzuki SO, Araki Y, Kuga D, Nakamizo A, Amano T *et al*: **MiRNA-196 Is Upregulated in**

Glioblastoma But Not in Anaplastic Astrocytoma and Has Prognostic Significance. *Clinical Cancer Research* 2010, **16**(16):4289-4297.
89. Guarino M, Rubino B, Ballabio G: **The role of epithelial-mesenchymal transition in cancer pathology**. *Pathology* 2007, **39**(3):305-318.
90. Legg J, Herbert J, Clissold P, Bicknell R: **Slits and Roundabouts in cancer, tumour angiogenesis and endothelial cell migration**. *Angiogenesis* 2008, **11**(1):13-21.
91. Khusial PR, Vadla B, Krishnan H, Ramlall TF, Shen Y, Ichikawa H, Geng J-G, Goldberg GS: **Src activates Abl to augment Robo1 expression in order to promote tumor cell migration**. *Oncotarget* 2010, **1**(3):198-209.
92. Humtsoe JO, Koya E, Pham E, Aramoto T, Zuo J, Ishikawa T, Kramer RH: **Transcriptional profiling identifies upregulated genes following induction of epithelial-mesenchymal transition in squamous carcinoma cells**. *Experimental Cell Research* 2012, **318**(4):379-390.
93. Labbé E, Lock L, Letamendia A, Gorska AE, Gryfe R, Gallinger S, Moses HL, Attisano L: **Transcriptional Cooperation between the Transforming Growth Factor-β and Wnt Pathways in Mammary and Intestinal Tumorigenesis**. *Cancer Research* 2007, **67**(1):75-84.
94. Hussain S-A, Piper M, Fukuhara N, Strochlic L, Cho G, Howitt JA, Ahmed Y, Powell AK, Turnbull JE, Holt CE *et al*: **A Molecular Mechanism for the Heparan Sulfate Dependence of Slit-Robo Signaling**. *Journal of Biological Chemistry* 2006, **281**(51):39693-39698.
95. Larrieu-Lahargue F, Welm AL, Thomas KR, Li DY: **Netrin-4 induces lymphangiogenesis in vivo**. *Blood* 2010, **115**(26):5418-5426.
96. Bartholow T, Chandran U, Becich M, Parwani A: **Immunohistochemical staining of radixin and moesin in prostatic adenocarcinoma**. *BMC Clinical Pathology* 2011, **11**(1):1.
97. Fuxe J, Vincent T, Garcia de Herreros A: **Transcriptional crosstalk between TGF beta and stem cell pathways in tumor cell invasion Role of EMT promoting Smad complexes**. *Cell Cycle* 2010, **9**(12):2363-2374.
98. Massague J: **How cells read TGF-beta signals**. *Nature Reviews Molecular Cell Biology* 2000, **1**(3):169-178.

Figures

Figure 10 – Overall analysis workflow

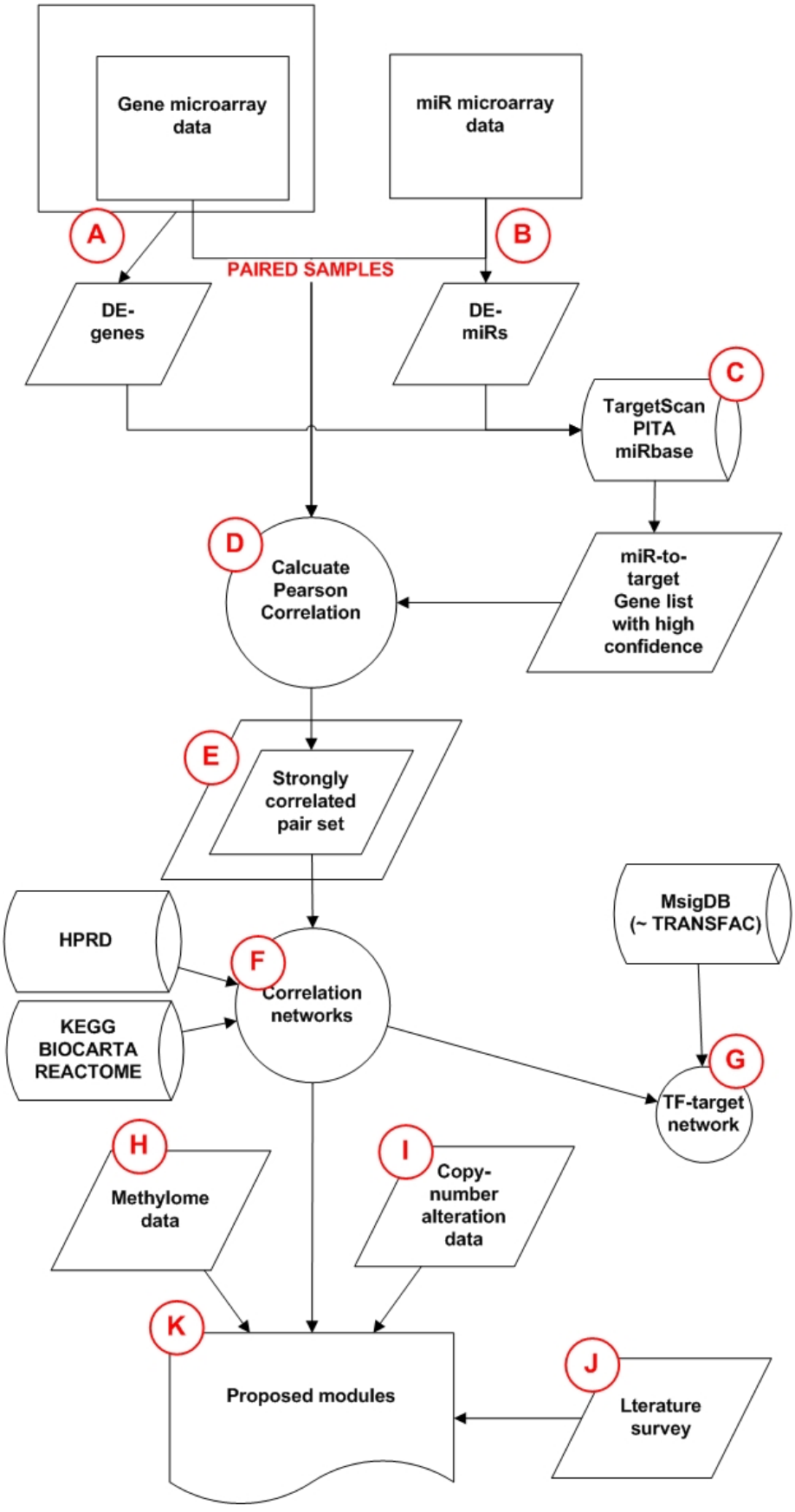


Figure 11 – Strong negative correlation network

Up-regulated miRNAs (diamond) and genes (circle) are represented in yellow or pink colors. Down-regulated miRNAs (diamond) and genes (circle) are represented in green or sky blue colors. Cancer-related genes (from Bushman Lab, see Methods section) are highlighted with yellow (up-regulated) or green (down-regulated). The gray edges between a miRNA and a gene represents a strong (thin, $0.4<|r|<0.75$) or a very strong (thick, $|r|>0.75$) Pearson correlation. Red edges between two genes represent their association in any canonical pathways defined in KEGG, BioCarta, or Reactome databases. Blue edges represent direct physical interaction between two genes defined in HPRD database. The size of each node represents the node's fold change in MPNST compared to NHSC.

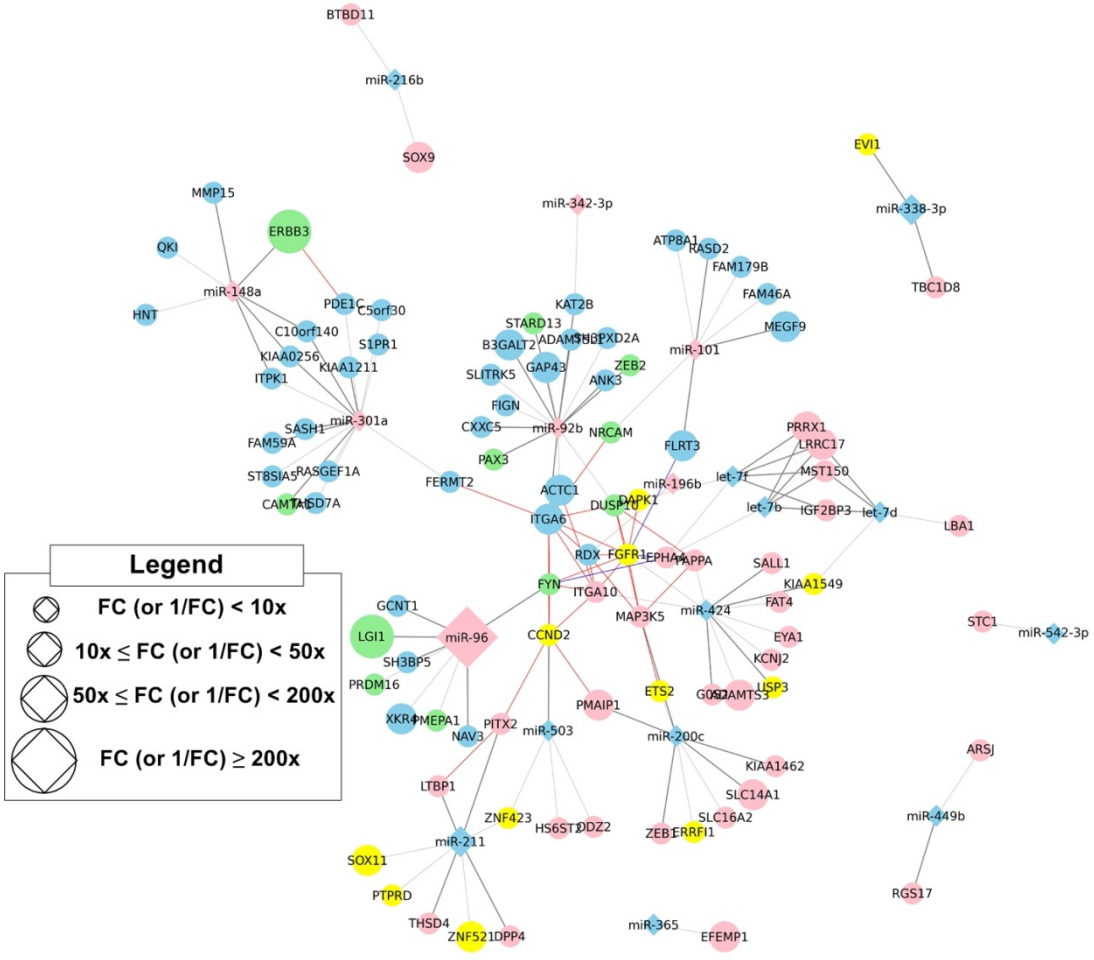


Figure 12 - Strong positive correlation networks

Up-regulated miRNAs (diamond) and genes (circle) are represented in yellow or pink colors. Down-regulated miRNAs (diamond) and genes (circle) are represented in green or sky blue colors. Cancer-related genes (from Bushman Lab, see Methods section) are highlighted with yellow (up-regulated) or green (down-regulated). The gray edges between a miRNA and a gene represents a strong (thin, $0.4<|r|<0.75$) or a very strong (thick, $|r|>0.75$) Pearson correlation. Red edges between two genes represent their association in any canonical pathways defined in KEGG, BioCarta, or Reactome databases. Blue edges represent direct physical interaction between two genes defined in HPRD database. The size of each node represents the node's fold change in MPNST compared to NHSC.

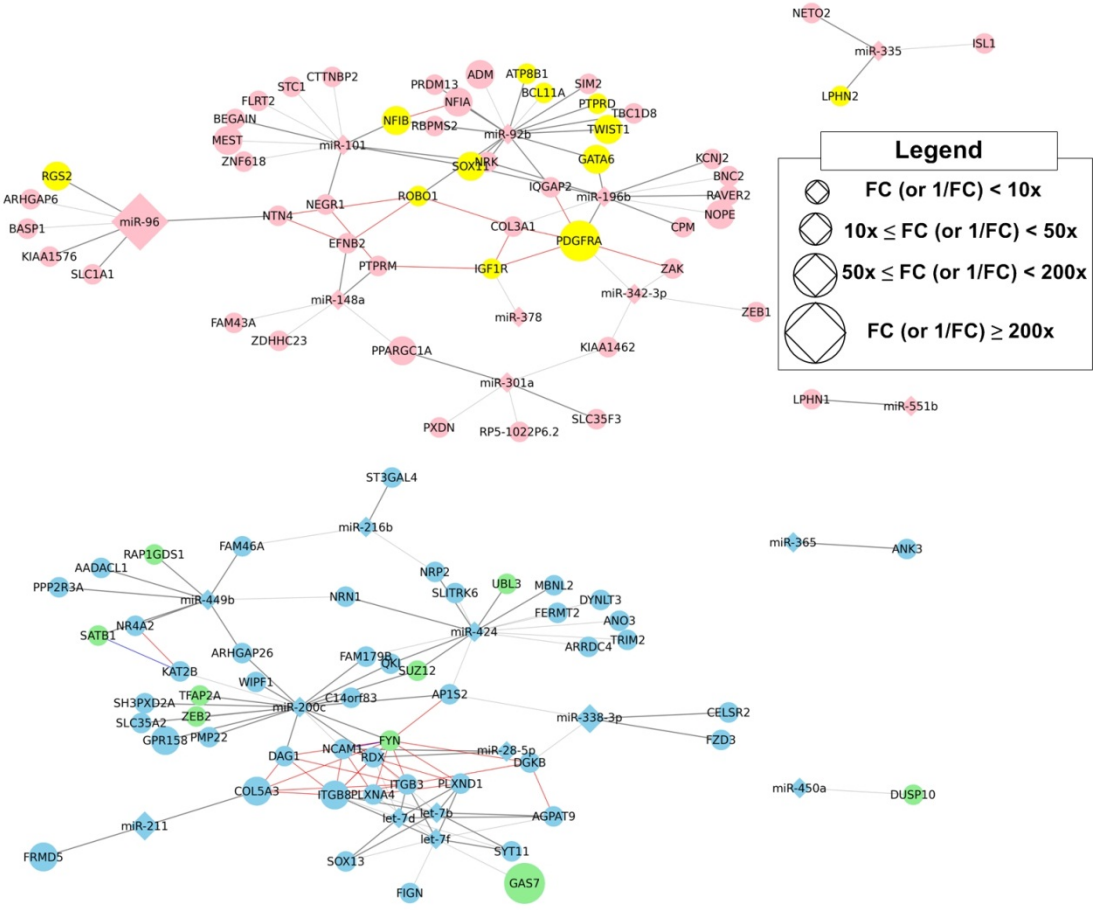


Figure 13 - Combined network

Figure 14 and Figure 15 are combined.

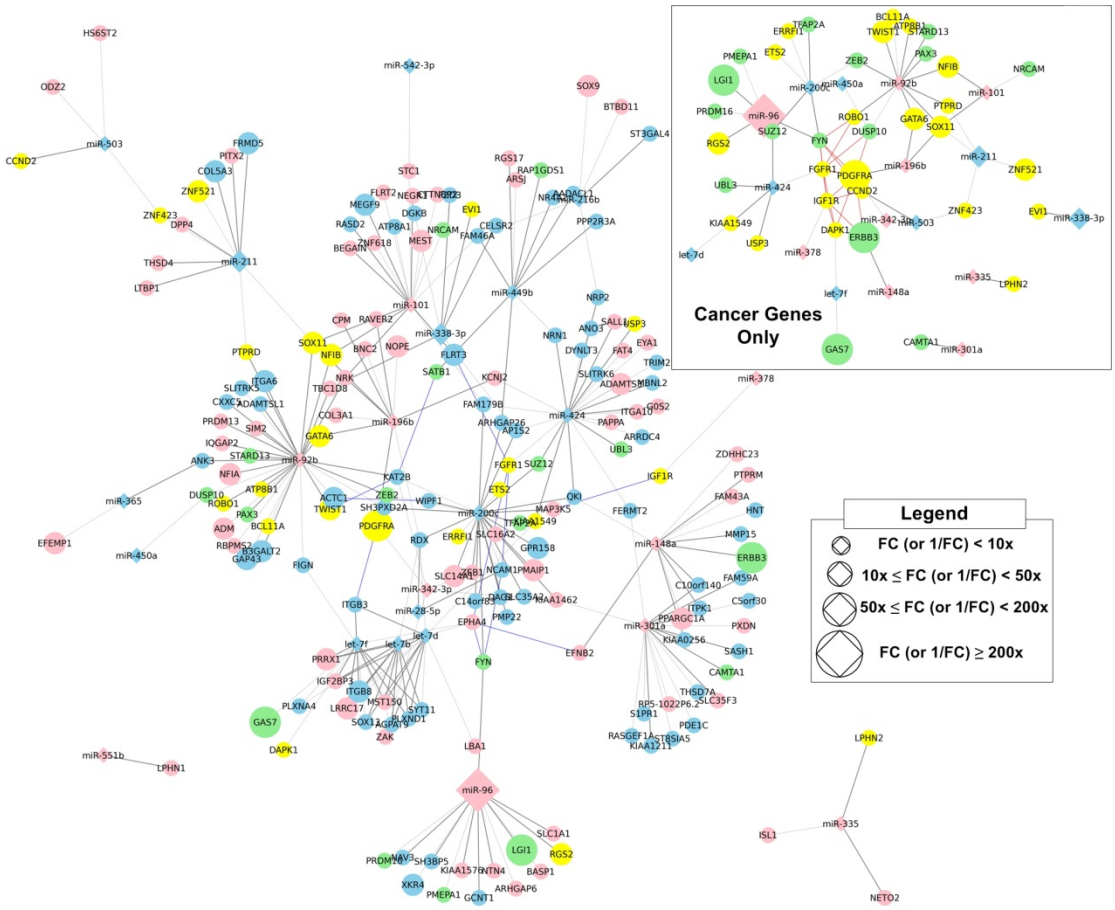


Figure 5 - TF-to-target gene network (A) and a proposed unit (B)

Pink/yellow nodes Blue/green nodes respectively represent up- and down-regulated miRNAs or genes. TFs are highlight using yellow or green colors. Four edge colors in panel A represent (i) thick gray (=strong correlation); (ii) thin gray (=stronger correlation); (iii) blue (=pathway membership); (iv) red (=experimentally proven protein-protein interaction); and (v) orange (=TF-target relationship).

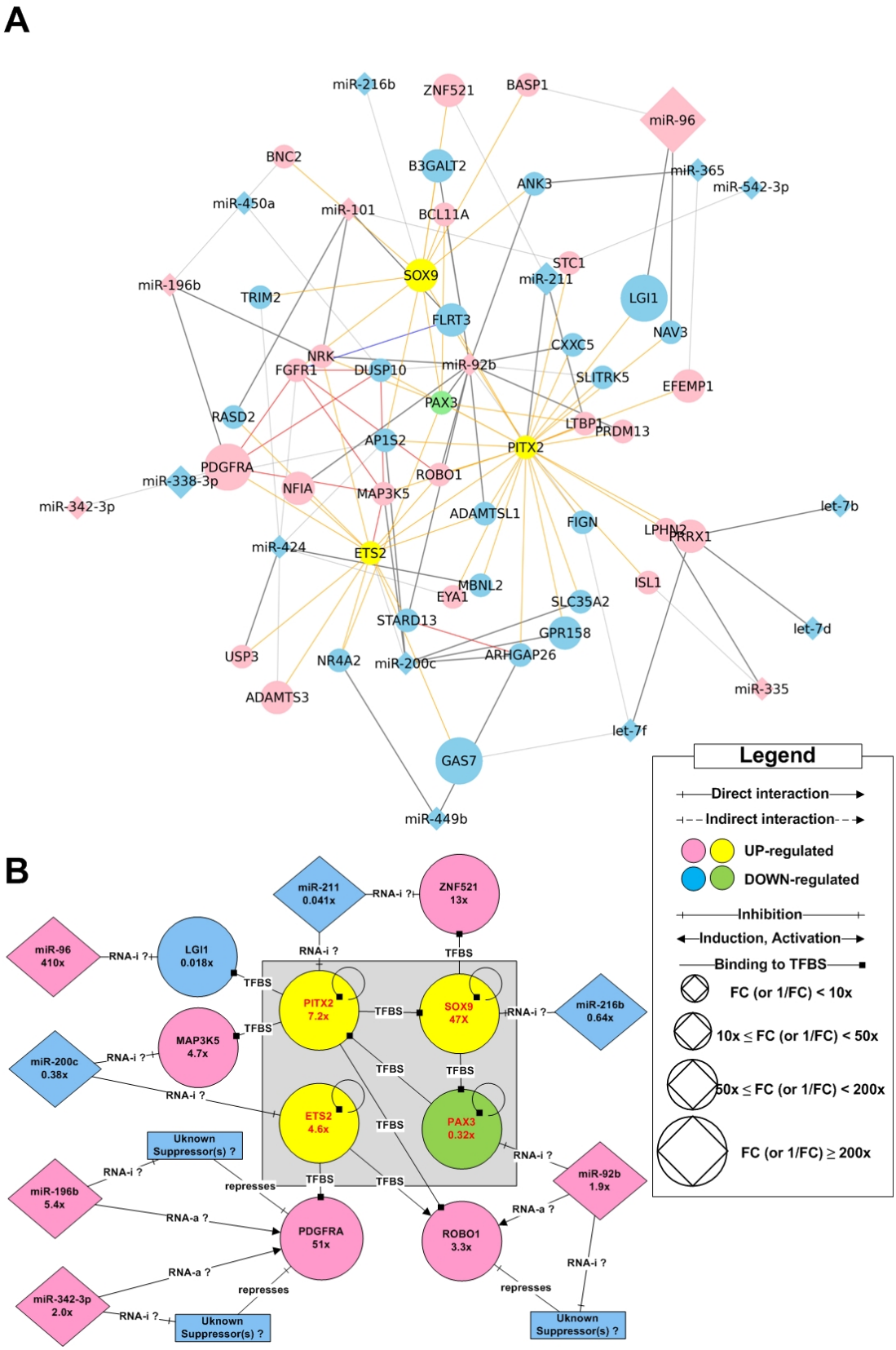

A
miR-216b
ZNF521
BASP1
miR-96
BNC2
B3GALT2
miR-365
ANK3
miR-542-3p
miR-450a
miR-101
BCL11A
STC1
SOX9
miR-211
miR-196b
TRIM2
LGI1
FLRT3
NAV3
CXXC5
NRK
FGFR1
DUSP10
miR-92b
SLITRK5
EFEMP1
PAX3
RASD2
LTBP1
PRDM13
AP1S2
PITX2
PDGFRA
ROBO1
miR-338-3p
NFIA
MAP3K5
miR-342-3p
ADAMTSL1
FIGN
let-7b
LPHN2
PRRX1
miR-424
ETS2
MBNL2
ISL1
EYA1
SLC35A2
STARD13
GPR158
let-7d
USP3
NR4A2
miR-200c
ARHGAP26
miR-335
ADAMTS3
let-7f
GAS7
miR-449b
Legend
Direct interaction
Indirect interaction
UP-regulated
DOWN-regulated
Inhibition
Induction, Activation
Binding to TFBS
FC (or 1/FC) < 10x
10x ≤ FC (or 1/FC) < 50x
50x ≤ FC (or 1/FC) < 200x
FC (or 1/FC) ≥ 200x
B
miR-211
0.041x
RNA-i ?
ZNF521
13x
TFBS
miR-96
410x
LGI1
0.018x
PITX2
7.2x
SOX9
47X
miR-216b
0.64x
MAP3K5
4.7x
miR-200c
0.38x
ETS2
4.6x
PAX3
0.32x
Uknown
Suppressor(s) ?
miR-196b
5.4x
miR-92b
1.9x
PDGFRA
51x
ROBO1
3.3x
RNA-a ?
represses
miR-342-3p
2.0x
Uknown
Suppressor(s) ?

Figure 6 - DE-genes in correlation networks and their relation with CNA and DNA methylation data.

DE-gene (up (=red)/down(=blue)) in the combined network are plotted in 3-D space. X-, Y-, and Z-axes represent differential DNA methylation (hypo-/hyper), fold change compared to NHSC, and copy number alteration patterns (gain/loss). Some interesting DE-genes which are probably controlled only by miR (yellow shade in the center) are listed on the side (green).

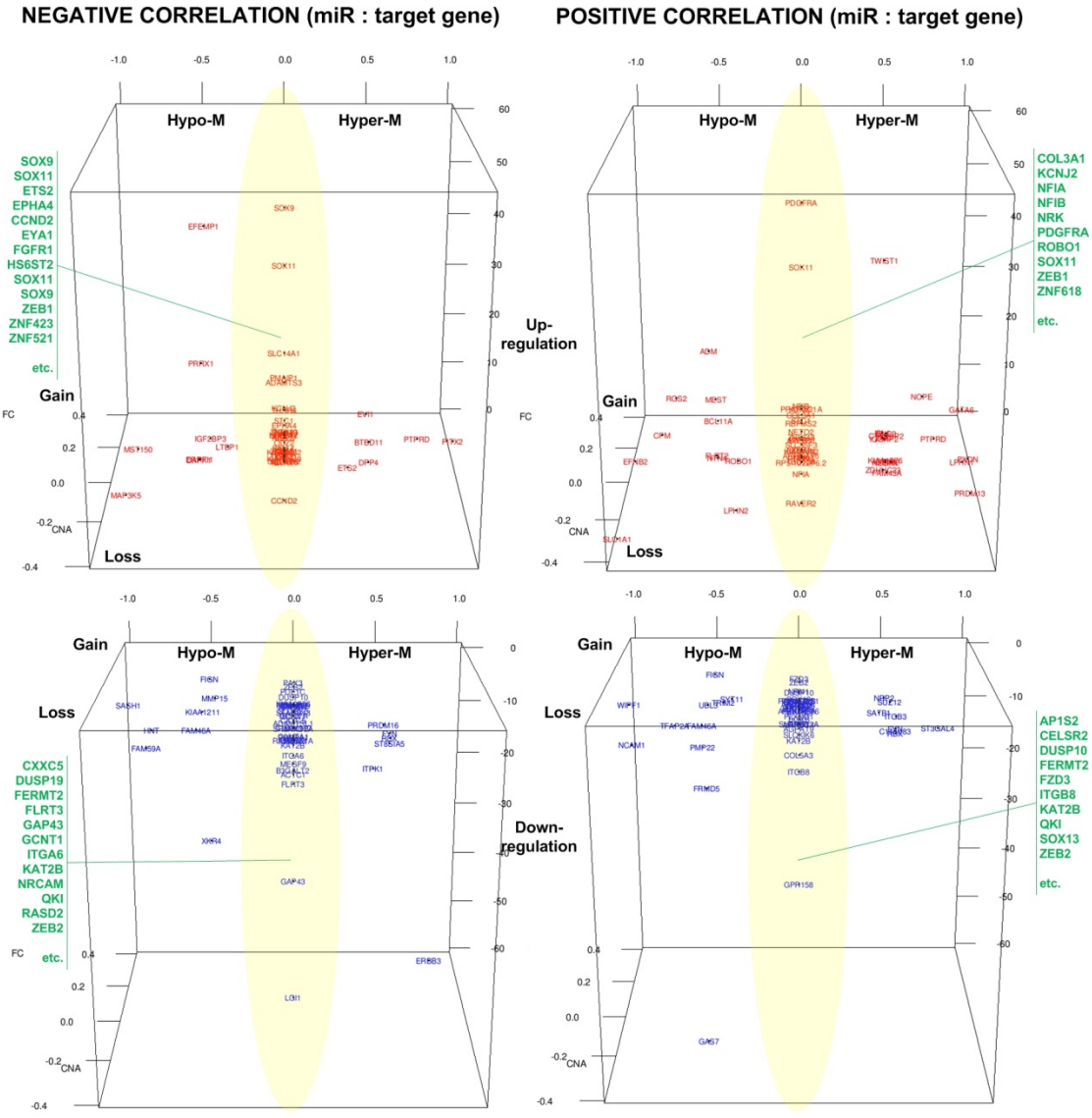


Figure 7 - Key modules reconstructed from correlation networks and literature search

Up-regulated miRNAs (diamond) and genes (circle) are represented as pink. Down-regulated miRNAs (diamond) and genes (circle) are represented as sky blue. If a

genes is not in correlation networks, but mentioned in Discussion section, the gene name is wrapped in parenthesis.

Tables

Table 1 - Network complexity before and after applying Pearson correlation filter

	Negative correlation						Positive correlation					
Filtering using Pearson correlation	**Nodes**		**Edges**	**Nodes**		**Edges**	**Nodes**		**Edges**	**Nodes**		**Edges**
	miR	**Gene**		**miR**	**gene**		**miR**	**gene**		**miR**	**Gene**	
	UP	**DN**		**DN**	**UP**		**UP**	**UP**		**DN**	**DN**	
Before	24	284	651	44	335	963	25	280	702	43	318	853
After	7	49	53	12	45	56	10	52	61	12	51	77

Table 2 – Summary of suggested regulatory modules

Models	miRNAs	Genes	Suggested function
A	miR-200c	ZEB1 TGFB2	TGF-beta induced signaling and reciprocal feedback loop [31, 32, 35, 41-43]
B	miR-338-3p miR-503 miR-211	EVI1 PITX2 ZNF43 ZNF521 (SMAD3) SOX10	Cell survival during EMT + TGF-beta/BMP/WNT pathway [46, 47, 53]
C	miR-101 miR-200c let-7d let-7f miR-424 miR-338-3p	FLRT3 NCAM1 FYN FGFR1 EPHA4 CELSR2 FZD3 (SOX10)	WNT/PCP pathway + Stemness control [11, 60-62, 64-68]
D	miR-92b miR-342-3p miR-200c	KAT2B TWIST1 PDGFRA (HMGA2)	Cell survival and ECM degradation [74, 75]
E	miR-196 miR-96	GATA6 LGI1 PRDM16 RDX PDGFRA (MMP1) (MMP3)	ECM degradation [36, 81-85, 87-89, 96]
F	miR-92b miR-96 miR-503	ROBO1 NTN4 HS5ST2	Cell migration [90, 93, 95]

Additional files

Additional file 1 – Genes, miRNAs, CNA DNA methylation data

Microsoft Excel file.

Additional file 2 –3D plot data (?)

Microsoft Excel file.

neg-corr-dn-up			
	FC	CNA	DM
ADAMTS3	18.37	0	0
ARSJ	4.164	0	0
BTBD11	7.077	0	0.4995
CCND2	3.724	-0.21429	0
DAPK1	3.584	0	-0.5025
DPP4	3.11	0	0.4995
EFEMP1	38.75	0.357143	-0.4995
EPHA4	3.06	0.214286	0
ERRFI1	3.65	0	-0.4995
ETS2	4.607	-0.07143	0.368
EVI1	7.715	0.142857	0.4995
EYA1	3.116	0	0
FAT4	3.942	0.142857	0
FGFR1	3.221	0	0
G0S2	5.236	0	0
HS6ST2	8.455	0	0
IGF2BP3	7.773	0	-0.4315
ITGA10	4.204	0.142857	0
KCNJ2	6.658	0.214286	0
KIAA1462	5.123	0	0
KIAA1549	3.859	0	0
LBA1	4.251	0	0
LRRC17	10.84	-0.07143	0
LTBP1	6.075	0	-0.3335
MAP3K5	4.686	-0.21429	-0.879
MST150	5.696	0	-0.8575
ODZ2	6.884	0	0
PAPPA	4.618	0	0
PITX2	7.199	0	0.994
PMAIP1	19.27	0	0
PRRX1	17.78	0.142857	-0.4995
PTPRD	7.625	0	0.7825
RGS17	7.836	0	0
SALL1	5.939	0	0
SLC14A1	23.85	0	0
SLC16A2	3.289	0	0
SOX11	32.87	0.285714	0
SOX9	46.89	0.142857	0
STC1	8.865	0.071429	0
TBC1D8	3.164	0	0
THSD4	6.172	0.214286	0
USP3	3.1	0.142857	0
ZEB1	5.499	0	0
ZNF423	6.605	0.071429	0
ZNF521	12.85	-0.21429	0

neg-corr-up-dn			
	FC	CNA	DM
ACTC1	-15.015	0	0
ADAMTSL1	-6.17284	0	0
ANK3	-7.04225	0	0
ATP8A1	-3.28947	0	0
B3GALT2	-14.3062	0	0
C10orf140	-6.84932	0	0
C5orf30	-3.03951	0	0
CAMTA1	-5.43478	-0.14286	0
CXXC5	-5.61798	0	0
DUSP10	-3.003	0.071429	0
ERBB3	-53.7634	0.142857	0.8295
FAM179B	-3.07692	0	0
FAM46A	-3.0303	-0.21429	-0.4995
FAM59A	-4.42478	-0.28571	-0.7435
FERMT2	-3.1746	0	0
FIGN	-3.63636	0.285714	-0.4995
FLRT3	-15.0376	-0.07143	0
FYN	-3.47222	-0.21429	0.4995
GAP43	-32.1543	-0.07143	0
GCNT1	-4.90196	0	0
HNT	-3.06748	-0.21429	-0.731
ITGA6	-11.7233	0	0
ITPK1	-9.17431	-0.21429	0.4175
KAT2B	-9.90099	0	0
KIAA0256	-4.14938	0.071429	0
KIAA1211	-4.21941	0	-0.4995
LGI1	-56.1798	0	0
MEGF9	-10	-0.14286	0
MMP15	-3.24675	0.071429	-0.44
NAV3	-3.02115	0	0
NRCAM	-9.17431	0	0
PAX3	-3.11526	0.214286	0
PDE1C	-4.62963	0.214286	0
PMEPA1	-8.54701	0	0
PRDM16	-6.49351	0	0.4995
QKI	-8.92857	0	0
RASD2	-9.00901	0	0
RASGEF1A	-4.69484	-0.21429	0
RDX	-4.31034	-0.21429	0.4995
S1PR1	-3.25733	0	0
SASH1	-3.09598	0	-0.9085
SH3BP5	-6.17284	0	0
SH3PXD2A	-7.0922	0	0
SLITRK5	-4.04858	0	0
ST8SIA5	-5.15464	-0.21429	0.4995
STARD13	-4.29185	0	0
THSD7A	-4.85437	0	0
XKR4	-34.2466	0.285714	-0.507
ZEB2	-3.83142	0.214286	0

pos-corr-up-up			
	FC	CNA	DM
ADM	30.25	-0.21429	-0.4995
ARHGAP6	4.308	0	0
ATP8B1	7.392	0	0
BASP1	5.041	0.071429	0
BCL11A	4.26	0.214286	-0.519
BEGAIN	5.268	0	0
BNC2	8.61	0	0.4995
COL3A1	3.212	0.285714	0
CPM	3.574	0.142857	-0.856
CTTNBP2	8.218	0	0.5
EFNB2	3.223	0	-0.984
FAM43A	3.513	-0.07143	0.4995
FLRT2	4.377	0	-0.4995
GATA6	19.99	-0.21429	0.88
IGF1R	3.186	0	0.4995
IQGAP2	5.251	0.071429	0.4995
ISL1	4.075	0.142857	0.4995
KCNJ2	6.658	0.214286	0
KIAA1462	5.123	0	0
KIAA1576	3.454	0	0.4995
LPHN1	3.073	0	0.955
LPHN2	4.152	-0.28571	-0.357
MEST	15.18	0	-0.4995
NEGR1	3.009	0	0.4995
NETO2	6.711	0.071429	0
NFIA	10.75	-0.28571	0
NFIB	13.9	0	0
NOPE	15.79	0	0.704
NRK	4.059	0	0
NTN4	3.896	0	-0.498
PDGFRA	50.82	0	0
PPARGC1A	13.31	0	0
PRDM13	4.656	-0.21429	0.9465
PTPRD	7.625	0	0.7825
PTPRM	3.77	0	0
PXDN	3.701	0	0.9885
RAVER2	5.412	-0.28571	0
RBPMS2	8.36	0.071429	0
RGS2	10.99	0.142857	-0.76
ROBO1	3.299	0	-0.3735
RP5-1022P	3.016	0	0
SIM2	5.525	-0.07143	0.5195
SLC1A1	4.378	-0.42857	-0.969
SLC35F3	6.503	0	0
SOX11	32.87	0.285714	0
STC1	8.865	0.071429	0
TBC1D8	3.164	0	0
TWIST1	35.79	0.214286	0.4995
ZAK	3.103	0.142857	0.4995
ZDHHC23	4.062	-0.07143	0.477
ZEB1	5.499	0	0
ZNF618	7.912	0	0

pos-corr-dn-dn			
	FC	CNA	DM
AADACL1	-5.07614	0.071429	0
AGPAT9	-7.87402	0	0
ANK3	-7.04225	0	0
ANO3	-3.09598	0	0
AP1S2	-5.07614	0	0
ARHGAP26	-4.85437	0	0
ARRDC4	-7.14286	0	0
C14orf83	-3.92157	-0.21429	0.4995
CELSR2	-4.56621	0	0
COL5A3	-15.456	0.142857	0
DAG1	-6.99301	0	0
DGKB	-6.21118	0	0
DUSP10	-3.003	0.071429	0
DYNLT3	-3.32226	0	0
FAM179B	-3.07692	0	0
FAM46A	-3.0303	-0.21429	-0.4995
FERMT2	-3.1746	0	0
FIGN	-3.63636	0.285714	-0.4995
FRMD5	-18.2149	0	-0.516
FYN	-3.47222	-0.21429	0.4995
FZD3	-3.1348	0.214286	0
GAS7	-57.4713	-0.21429	-0.4995
GPR158	-27.3224	-0.28571	0
ITGB3	-5.74713	0	0.5365
ITGB8	-20.202	0.214286	0
KAT2B	-9.90099	0	0
MBNL2	-4	0	0
NCAM1	-6.13497	-0.21429	-0.8665
NR4A2	-3.95257	0	0
NRN1	-4.13223	0.142857	0
NRP2	-6.57895	0.214286	0.4995
PLXNA4	-3.64964	0	0
PLXND1	-5.74713	0	0
PMP22	-6.45161	-0.21429	-0.4995
PPP2R3A	-4.78469	0	0
QKI	-8.92857	0	0
RAP1GDS1	-3.04878	0	0
RDX	-4.31034	-0.21429	0.4995
SATB1	-5.10204	0	0.442
SH3PXD2A	-7.0922	0	0
SLC35A2	-4.29185	0	0
SLITRK6	-8.92857	0	0
SOX13	-5.43478	0.142857	0
ST3GAL4	-3.42466	-0.21429	0.7175
SUZ12	-3.27869	0	0.4995
SYT11	-4.16667	0.071429	-0.377
TFAP2A	-7.35294	0	-0.6875
TRIM2	-3.367	0	-0.416
UBL3	-3.7594	0	-0.4995
WIPF1	-3.7594	0	-0.937
ZEB2	-3.83142	0.214286	0

Additional file 3 –TF-and-target network

Microsoft Excel file.

BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Database

Volume 8(15)

Competing endogenous RNA database

Aaron L Sarver[1] & Subbaya Subramanian[1, 2*]

[1]Masonic Cancer Center; [2]Division of Basic and Translational Research, Department of Surgery, University of Minnesota, MN 55455; Subbaya Subramanian – E mail: subree@umn.edu; Phone: 612-626-4330; Fax: 612-626-7031; *Corresponding author

Received July 27, 2012; Accepted July 28, 2012; Published August 03, 2012

Abstract:
A given mRNA can be regulated by interactions with miRNAs and in turn the availability of these miRNAs can be regulated by their interactions with alternate mRNAs. The concept of regulation of a given mRNA by alternate mRNA (competing endogenous mRNA) by virtue of interactions with miRNAs through shared miRNA response elements is poised to become a fundamental genetic regulatory mechanism. The molecular basis of the mRNA-mRNA cross talks is via miRNA response elements, which can be predicted based on both molecular interaction and evolutionary conservation. By examining the co-occurrence of miRNA response elements in the mRNAs on a genome-wide basis we predict competing endogenous RNA for specific mRNAs targeted by miRNAs. Comparison of the mRNAs predicted to regulate PTEN with recently published work, indicate that the results presented within the competing endogenous RNA database (*ceRDB*) have biological relevance.

Availability: http://www.oncomir.umn.edu/cefinder/

Key words: ceRNAs, MRE, microRNA response elements, database, competing endogenous RNAs database, ceRDB

Background:
MicroRNAs (miRNAs) play an important role in almost all biological functions **[1].** Transcriptional deregulations in miRNAs have been implicated in disease processes including cancers and developmental disorders **[2]**. It has been well established that a single miRNA can regulate the expression of many mRNAs/ transcripts and an mRNA can be regulated by multiple miRNAs **[1]**. miRNA gene regulation is mediated by a complex set of proteins termed RNA induced silencing complex. The miRNAs are guided to the miRNA response elements (MRE) present in the target mRNAs, which may result in transcript degradation and/or translational inhibition **[3].** Recently it has been established that miRNA activity on the target gene can be influenced by the presence or absence of other competing endogenous (ceRNA) mRNAs that contain shared MREs **[4-7]**. These miRNA activity modulators can act as a sponge, absorbing and releasing miRNA based on the level of the transcript. Several modulators of miRNA activity have been recently characterized **[8].** Salmena et al proposed a hypothesis that these modulators can communicate with each other in a miRNA dependent manner mediated through MREs **[9].** This complex miRNA-mRNA network and interactions opens up a new chapter in miRNA-mediated gene regulation. However, currently there are no publicly available resources that identify and catalog the list of genes that can act as miRNA activity modulators or ceRNAs. Here we developed a comprehensive and easy to use resource named '*competing endogenous RNA database (ceRDB)*' that lists potential MRE containing genes that can act in a sponge like fashion for a given mRNA based on a set of scoring and ranking criteria.

Methodology:
MiRNA-mRNA target interactions were obtained from http://www.targetscan.org Release 5.2 June 2011. The predicted conserved target information file was parsed to obtain 54979 conserved human miRNA-mRNA interactions. To explore the structure of the dataset, the list of interactions was converted into a matrix containing 153 miRNA families on the X-axis and 9448 target mRNAs on the Y-axis. The presence of a predicted conserved miRNA-mRNA interaction is defined by the presence of a '1' at the defined gene row miRNA column corresponding to the interaction. The absence of an interaction is defined by the presence of a '0' at the corresponding interaction. To shuffle the matrix, interactions between each miRNA and mRNA were randomly assigned maintaining the total number of interactions for each mRNA. Both the real matrix and the shuffled matrix were filtered to only show genes with more than 5 miRNA binding sites and these were

clustered using Gene Cluster 3.0 hierarchical clustering of both the X and the Y-axis using Centroid linkage. The resulting clustered matrixes were visualized using Java Treeview. To score potential ceRNA interactions, the 54979 human interactions were loaded into a mySQL database and when the user selects a given mRNA all predicted miRNA targets for the given mRNA are obtained. These miRNA are then used to define all mRNAs that contain binding sites for the set of miRNAs. For each mRNA, an interaction score is then defined by adding up the total number of miRNA binding sites that overlap with the miRNA for a given mRNA. This interaction score is then used to sort the results and the top 50 predicted potential ceRNAs are returned. This process is carried out on the fly using PHP interactions with mySQL in a similar fashion as previously described in our publicly available databases such as sarcoma microRNA expression database (S-MED).

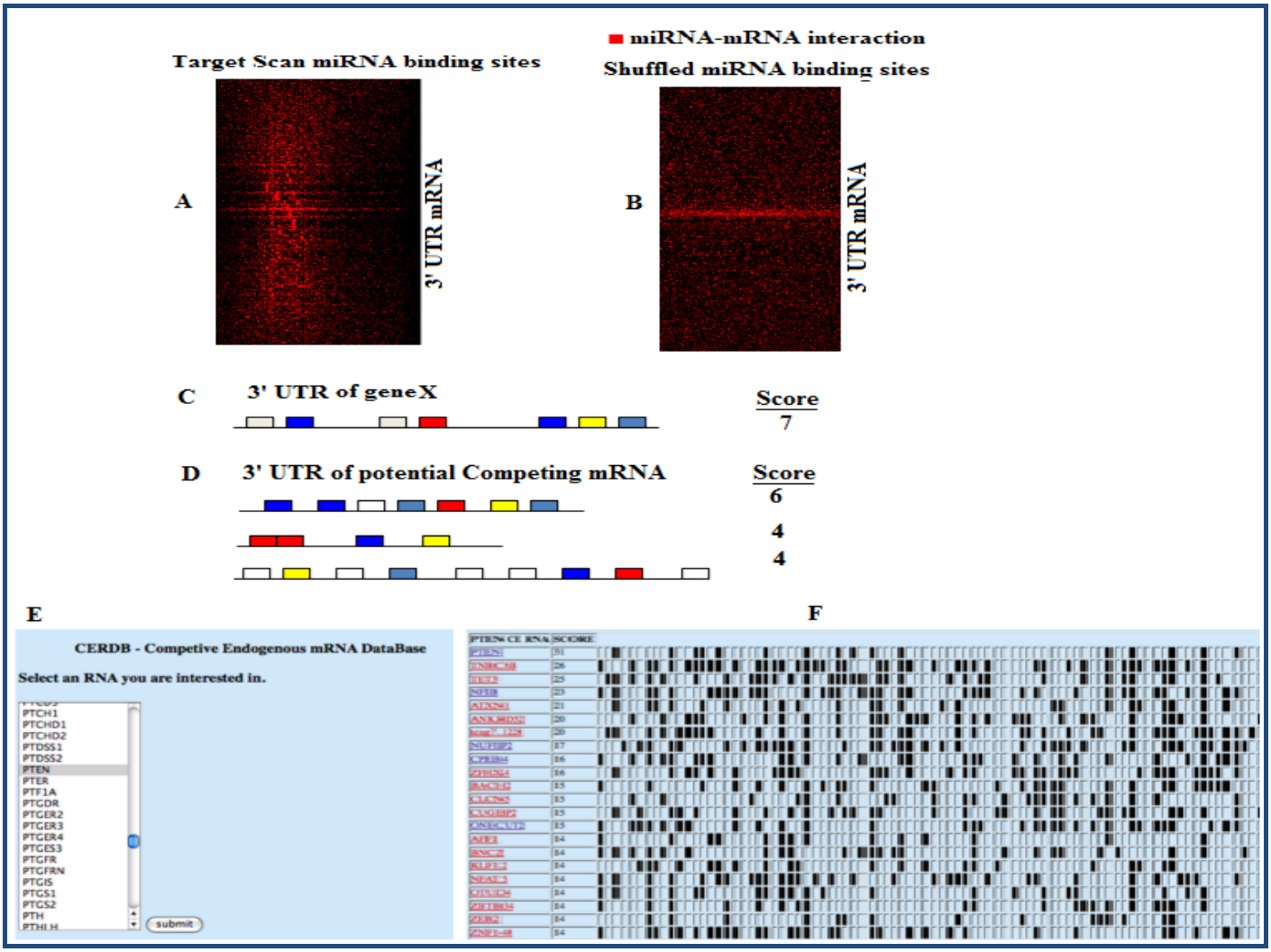


Figure 1: Visualization of co-occurrence in predicted miRNA-mRNA interactions**.** Heat map showing the presence of predicted miRNA-binding sites on the X-axis and the genes that contain the binding sites in the 3'UTR on the Y-axis. Only genes that show more than 5 binding sites are shown for **(A)** predicted interactions and **(B)** predicted interactions after shuffling. **(C)** Predicting competing mRNA via miRNA-mRNA interactions. miRNA binding site predictions in the 3'UTR are shown as colored boxes. The 'Score' is generated by counting the number of conserved predicted interactions. In this hypothetical case shown there are 7 predicted binding elements in the 3'UTR of the gene. **(D)** To predict potential competing RNA for the gene shown in A, binding sites for the predicted miRNA found in A are obtained and summed in all genes. The genes are then sorted by total number of overlapping binding sites and returned to the user. **(E)** Example of competing mRNA predictions from ceRDB for PTEN. The user selects an mRNA of interest from the list of available mRNA. In the case shown here the PTEN tumor suppressor is chosen. **(F)** Starting with the list of miRNA binding elements present in PTEN the tool predicts potential competing RNA and visualizes the extent of overlap between the miRNA binding sites. Only a representative subset of the matrix is shown, the full matrix is available online. Each predicted gene is linked back to the TargetScan database to visualize the position and total numbers of each miRNA element.

Results:

In order to define the information content present within miRNA-mRNA predicted interactions we clustered a matrix containing miRNA families on the Y-axis with genes on the X-axis. Predicted binding interactions are labeled with a '1' and the lack of an interaction is labeled with a '0' at corresponding

points in the matrix. A heatmap of the clustered images as well as the branch structure indicate that miRNA binding sites coexist within 3′UTR at a much higher rate than would be expected at random. To visually show this we randomized the interaction matrix and clustered the results **(Figure 1A & B).** Within the cell, each miRNA has many mRNA targets and each mRNA has potentially many miRNAs capable of regulation leading to a complex and dynamic regulatory system. One heretofore overlooked consequence of this system is that manipulation of the transcript level of a given mRNA may lead to changes in the concentration of available miRNAs leading to changes in alternate mRNA regulation via miRNA-mRNA interactions. In order to predict these interactions for a given target mRNA we determined all possible miRNA binding to this target mRNA and then found mRNAs (ceRNAs) that contained binding sites to these miRNAs. The potential for competition was ranked for each mRNA by counting the number of overlapping miRNA binding sites shared between the given mRNA and the potential ceRNA **(Figure 1C & D).** Competing endogenous mRNA rankings were generated using the conserved mRNA-miRNA interactions. To access the data, we built a simple to use web interface and have made it available at http://www.oncomir.umn.edu/cefinder/. The user enters an mRNA they are interested in finding potential competing mRNAs that can regulate the gene of interest, and the tool returns potential ceRNA regulators. The list is sorted based on the overlap of the miRNA binding sites in the each of the pairwise relationships **(Figure 1E & F).** Additionally the miRNA interactions present within the 3′UTR of the primary mRNA and all potential regulators are visualized in the final table**.**

Discussion:

In recent years miRNAs have taken center stage in many aspects of post transcriptional gene regulation. The complexity of miRNA-mediated gene regulation is compounded by the presence of multiple mechanisms that modulate either the levels of miRNAs and/or its target gene. Recently, the Pandolfi group proposed a novel concept in which mRNAs can regulate each other via common miRNA response elements **[4, 8]**. Through this cross talk novel mRNA-mRNA interactions have been identified in multiple cancer types. These findings suggest that modulation of miRNA activity by changing the levels of competing endogenous RNA is a key fundamental mechanism of gene regulation that will be applicable for many biological functions. Here we present a general and straightforward tool for identifying competing endogenous RNAs (ceRNAs) for a given gene of interest. Starting with the conserved set of miRNA-mRNA interactions, we observe that there is high degree of co-occurrence of miRNA binding sites within the miRNA-mRNA interaction dataset. This is consistent with the reports of Shalgi et al **[10].** We then use the co-occurrence of miRNA binding sites to predict and rank potential ceRNAs for all mRNAs. Our predictions are experimentally validated for PTEN and likely very relevant for a large number of additional genes **[5].** Several recent articles have described ceRNAs that are capable of regulating *PTEN* via competing reactions **[4, 5].** In these cases, loss of a competing mRNA releases miRNAs for interaction with the tumor suppressor *PTEN* leading to decreased *PTEN* expression. Our database predicts many of the biologically validated interactions previously reported and uses a very straightforward algorithm in identifying these competing endogenous RNAs. Our search for ceRNAs for many established tumor suppressors in our database revealed some interesting observations. For example, genes such as *ONECUT2*, *NFIB* and *TNRC6B* appeared in many of the ceRNAs gene lists, these genes contains long 3′UTRs of up to 14kb in length and are predicted to contain many MREs that can potentially act as a sponge for multiple miRNAs. We are tempted to speculate that these ceRNAs with long 3′UTR can act as a 'master' MRE containing gene whose regulation may be affected in multiple disease conditions. Recently, TNRC6B was predicted to function as a ceRNA for *PTEN* and the downregulation of *TNRC6B* reduced the expression of *PTEN* **[5].**

In conclusion, we have developed the ceRDB resource to in the future accommodate multiple species such as model organisms and other types of sequences such as long non-coding RNAs and pseudogenes that can potentially also function as ceRNAs. We believe that the concept of competing endogenous RNA is likely to become a canonical central theme of gene regulation and having the ceRDB resource will significantly enhance our understanding of this fundamental gene regulatory mechanism.

Conflict of Interest:

We declare no conflict of interests

Author contributions:

AS and SS developed the idea. AS wrote the code and implemented. AS and SS wrote the manuscript.

Funding:

We acknowledge the funding from Department of Defense grant number # W81XWH-10-1-0556.

Acknowledgement:

We profusely thank Drs. Clifford Steer, Reena Kartha, Praveensingh Hajeri and Venugopal Thayanithy for their helpful discussions. We also wish to acknowledge the Minnesota Supercomputing Institute for providing access to computational resources.

References:

[1] Bartel DP, *Cell.* 2009 **136**: 215 [PMID: 19167326]
[2] Van Kouwenhove M *et al. Nat Rev Cancer*. 2011 **11**: 644 [PMCID: 21822212]
[3] Guo H *et al. Nature.* 2010 **466**: 835 [PMID: 20703300]
[4] Tay Y *et al. Cell*. 2011 **147**: 344 [PMID: 22000013]
[5] Karreth FA *et al. Cell.* 2011 **147:** 382 [PMID: 22000016]
[6] Cesana M *et al. Cell.* 2011 **147**: 358 [PMID: 22000014]
[7] Sumazin P *et al. Cell.* 2011 **147**: 370 [PMID: 22000015]
[8] Poliseno L *et al. Nature.* 2010 **465**: 1033 [PMID: 20577206]
[9] Salmena L *et al. Cell.* 2011 **146**: 353 [PMID: 21802130]
[10] Shalgi R *et al. PLoS Comput Biol.* 2007 **3**: e131 [PMID: 17630826]

Edited by P Kangueane

Citation: Sarver & Subramanian, Bioinformation 8(15): 731-733 (2012)